Technische Universität München

Fakultät für Chemie

Lehrstuhl für biomolekulare NMR Spektroskopie

Structural and functional characterization of the *ASH1*-mRNP transport-complex from budding yeast

Franziska Theresia Edelmann

Vollständiger Abdruck der von der Fakultät für Chemie an der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Bernd Reif

Prüfer der Dissertation: 1. Univ.-Prof. Dr. Michael Sattler

                                                       2. Univ.-Prof. Dr. Franz Hagn

                                                       3. Univ.-Prof. Dr. Dierk Niessing

Die Dissertation wurde am 14.02.2017 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 28.03.2017 angenommen.

.

Abstract

mRNA localization is an important mechanism of gene regulation and is required to modulate cellular architecture and function. Dedicated RNA-binding proteins (RBPs) selectively recognize *cis*-acting localization elements (LEs) in mRNAs for specific transcript localization. They form stable messenger ribonucleoprotein particles (mRNPs), which incorporate motor proteins to fulfill active cargo transport. Despite detailed analyses by various imaging approaches, the assembly- and processivity-mediating features of motile particles for asymmetric mRNA localization are not well understood. Additionally there is still a lack of knowledge about which features are responsible for LE specificity in an mRNP. A first step towards a better mechanistic understanding was achieved by the *in vitro* recapitulation of the highly stable and specific recognition of *ASH1* mRNA by the RBP She2p and the Myo4p motor-bound RBP She3p in budding yeast. Therefore budding yeast serves as perfect model system to obtain detailed mechanistic information on these events.

The first part of my study served the examination of mRNP assembly- and processivity-mediating features. Single-particle motility assays revealed that *in vitro* reconstituted SHE particles consisting of Myo4p, She2p, She3p and the *ASH1* mRNA indeed move processively along actin filaments. Observed run lengths are compatible with the long-distance transport reported *in vivo*. The results further showed that cargo RNA is dispensable for motor activation and that rather the protein-protein interaction between She2p and She3p is essential to activate motility. Stability analysis of the complex in absence of RNA proved that the observed binding events are robust and specific, which directly supports the findings of single-particle motility assays. Further analysis revealed that the reconstituted SHE complexes have a defined size and are able to multimerize into larger particles upon binding of RNAs with multiple LEs. However, this does not affect particle motility or run length.

The second part of my study provides the first comprehensive structural and functional analysis of how an mRNA-transport complex specifically recognizes its stem loop-containing LE RNA. My work recapitulates and characterizes the step-wise maturation process of the *ASH1* mRNP on a structural level, beginning with an LE in its unbound state, via the nuclear She2p-RNA co-complex, to the highly stable and specific cytoplasmic complex consisting of She2p, She3p and the RNA. A combination of X-ray crystallography, nuclear magnetic resonance (NMR) and small-angle X-ray scattering (SAXS) experiments showed that one of the *cis*-acting LEs of the *ASH1* mRNA alone adopts an elongated stem loop with a highly dynamic central bulged region between its flanking helices. Upon co-transcriptional binding She2p recognizes the RNA in its kinked state most likely via conformational selection from a dynamic ensemble of different RNA folding states. The transition from elongated to kinked state requires the RNA to undergo dramatic conformational changes. In the binary complex

three bases (C1779, U1780 and C1813) are specifically recognized by the RBP She2p. Additional joining of the unstructured She3p induces the formation of a highly specific, synergistic cytoplasmic complex. This increase in specificity and affinity in the mature transport complex is mainly achieved by a more precise recognition of the RNA shape via an unstructured region of She3p. Although a multi-interaction network of She2p and She3p recognizes a range of features in the LE RNA, the two conserved cytosines already recognized by She2p alone remain the only sequence-specifically recognized bases in the ternary complex. However, correct *in vivo ASH1*-mRNA localization also requires additional protein-protein interactions between She2p and She3p. My findings clarify that it is a combination of structural and sequence requirements that are necessary for RNA binding. These aspects and the dramatic rearrangements of the RNA upon binding could have not been predicted before and therefore offer a significant advancement in our understanding of specific mRNA recognition for its transport.

Zusammenfassung

mRNA-Lokalisation ist ein wichtiger Mechanismus der Genregulation und wird benötigt, um die zelluläre Architektur und Funktion zu modulieren. Bestimmte RNA-bindende Proteine (RBPs) erkennen gezielt *cis*-aktive Lokalisationselemente (LEs) in mRNAs zur spezifischen Lokalisation der Transkripte. Sie bilden stabile Messenger-Ribonukleoproteinpartikel (mRNP), welche Motorproteine für den aktiven Gütertransport enthalten. Trotz detaillierter Analysen mit vielfältigen mikroskopischen Ansätzen sind die Eigenschaften, welche den Zusammenbau und die Prozessivität der beweglichen Partikel vermitteln, nicht gut verstanden. Zudem fehlt das Wissen, welche Merkmale für die LE-Spezifität in einem mRNP verantwortlich sind. Ein erster Schritt zu einem besseren mechanistischen Verständnis war die *in vitro*-Rekapitulation der hochstabilen und spezifischen Erkennung der *ASH1*-mRNA durch das RBP She2p und das Myo4p-motorgebundene RBP She3p aus der Bäckerhefe. Daher dient die Bäckerhefe als perfektes Modellsystem, um ein detailliertes mechanistisches Verständnis dieser Ereignisse zu erhalten.

Der erste Teil meiner Arbeit diente der Untersuchung der Eigenschaften, die den Zusammenbau und die Prozessivität der mRNPs vermitteln. Motilitätsanalysen mit einzelnen Partikeln zeigten, dass sich *in vitro* nachgebildete SHE-Partikel, welche aus Myo4p, She2p, She3p und der *ASH1*-mRNA bestehen, tatsächlich aktiv auf Aktin *in vitro* bewegen. Die beobachteten Lauflängen sind dabei derer aus *in vivo*-Transportvorgängen sehr ähnlich. Die Ergebnisse zeigten weiter, dass die RNA-Fracht für die motorische Aktivierung entbehrlich ist und dass es vielmehr die Protein-Protein-Wechselwirkung zwischen She2p und She3p ist, welche die Aktivierung der Motilität vermittelt. Die Stabilitätsanalyse des Komplexes in Abwesenheit der RNA beweist, dass die beobachteten Bindungsereignisse robuste und spezifische Wechselwirkungen sind, welche die Ergebnisse der Motilitätsanalysen mit einzelnen Partikeln direkt stützen. Weitere Analysen zeigten, dass nachgebildete SHE Komplexe eine definierte Größe besitzen und dazu fähig sind sich in größere Partikel zusammenzulagern. Dies hat jedoch keinen Einfluss auf die Beweglichkeit des Partikels oder auf seine Lauflänge.

Der zweite Teil meiner Studie liefert die erste umfassende strukturelle und funktionelle Analyse, wie ein mRNA-Transport-Komplex spezifisch seine haarnadelförmige LE-RNA erkennt. In meiner Arbeit wird der stufenweise Reifungsprozess des *ASH1*-mRNP strukturbiologisch nachgebildet und charakterisiert, beginnend mit einem LE in ungebundenem Zustand über den nuklearen She2p-RNA-Kokomplex zum hochstabilen und spezifischen, zytoplasmatischen Komplex bestehend aus She2p, She3p und der RNA. Eine Kombination von Röntgenkristallographie-, Kernspinresonanz- (NMR-) und Röntgenkleinwinkelstreuungs- (SAXS-) Experimenten zeigte, dass eines der *cis*-wirkenden LE der *ASH1*-mRNA alleine eine langgestreckte Haarnadelschleife mit einem hochdynamischen zentralen,

aufgeweiteten Bereich ausbildet. Bei der kotranskriptionellen Bindung erkennt She2p die RNA im geknickten Zustand, höchstwahrscheinlich durch konformationelle Selektion aus einem dynamischen Ensemble verschiedener RNA-Faltungszustände. Der Übergang vom gestreckten zum geknickten Zustand erfordert, dass die RNA dramatische Konformationsänderungen erfährt. Im Binärkomplex werden drei Basen (C1779, U1780 und C1813) spezifisch vom RBP She2p erkannt. Eine zusätzliche Anlagerung des unstrukturierten She3p induziert die Bildung eines hochspezifischen, synergistischen, zytoplasmatischen Komplexes. Diese Erhöhung der Spezifität und Affinität in dem reifen Transportkomplex wird hauptsächlich durch eine präzisere Erkennung der Form der RNA über einen ungefalteten Bereich von She3p erreicht. Obwohl ein Multi-Interaktionsnetzwerk von She2p und She3p eine Reihe von Merkmalen in der LE-RNA erkennt, sind die beiden konservierten, auch von She2p alleine erkannten Cytosine die einzigen sequenzspezifisch erkannten Basen im ternären Komplex. Zusätzlich erfordert die korrekte *ASH1* mRNA-Lokalisierung *in vivo* auch weitere Protein-Protein-Interaktionen zwischen She2p und She3p. Meine Ergebnisse klären, dass es eine Kombination aus Struktur- und Sequenzanforderungen ist, die für die spezifische RNA-Bindung notwendig sind. Diese Aspekte und die dramatischen Umlagerungen der RNA bei der Bindung konnten vorher nicht vorhergesagt werden und leisten deshalb einen wichtigen Beitrag unser Verständnis der spezifischen mRNA-Erkennung für den Transport voranzutreiben.

Table of Contents

List of Figures

List of Tables

Abbreviations

2'-ACE	2'-O-bis-(2-acetoxyethoxy)methyl	NOESY	nuclear overhauser enhancement spectroscopy
A	absorption		
aa	amino acid	nt	nucleotides
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluoride	OD	optical density
		o.n.	over night
Amp	ampicillin	PAGE	polyacrylamide gel electrophoresis
APS	ammonium peroxydisulfate	PCR	polymerase chain reaction
a.u.	arbitary units	PEG	polyethylene glycol
ATP	adenosine-5'-triphosphate	PEPF	Protein expression and purification facility
bp	base pairs		
BSA	bovine serum albumine	PMSF	phenylmethylsulfonyl fluoride
Cam	chloramphenicol	PTFE	polytetrafluorethylene
CaM	calmodulin	PVDF	Polyvinylidenfluoride
CD	circular dichroism	RMSD	root-mean-square deviation
Cmd	calmodulin	RNA	ribonucleic acid
CV	column volume	rpm	revolutions per minute
Da	dalton	RT	room temperature
dd	double-destilled	*S. cerevisae*	*Saccharomyces cerevisiae*
DEPC	diethylpyrocarbonate	s.d.	standard deviation
DIC	differential interference contrast	SAXS	small-angle X-ray scattering
DMSO	dimethyl sulfoxide	SDS	sodium dodecyl sulfate
DNA	deoxyribonucleic acid	SEC	size exclusion
D. melanogaster	*Drosophila melanogaster*	*Sf*	*Spodoptera frugiperda*
		SHE	Swi5p-dependent HO expression
DTT	1,4-dithiothreitol	SOB	super optimal broth
E. coli	*Escherichia coli*	SOC	super optimal catabolite repression
EDTA	ethylendiamintetraacetic acid	TAE	Tris/acetate/ethylendiamintetraacetic acid
EGTA	ethylene glycol tetraacetic acid		
ER	endoplasmic reticulum	TCEP	Tris-(2-carboxyethyl)-phosphine
G6PD	Glucose-6-phosphate dehydrogenase	TEMED	N,N,N´,N´-Tetramethylethylendiamine
Gent	gentamicin	TMR	tetramethylrhodamine
GFP	green-fluorescent protein	Tet	tetracycline
GST	gluthatione S-transferase	TEV	tobacco etch virus
HMGU	Helmholtz-Zentrum München	TGS	Tris/glycine/sodium dodecyl sulfate
HPLC	high performance liquid chromatography	Thr	thrombin
		TIRFM	total internal reflection fluorescence microscopy
HRV	human rhinovirus		
IPTG	isopropyl-β-D-thiogalaktopyranosid	TL/TLR	tetraloop/ tetraloop receptor
ITC	isothermal titration calorimetry	TLS	translation-liberation-screw rotation
Kan	kanamycin	U	unit
LMW	low molecular weight	UV	ultraviolet
MBP	maltose-binding protein	v/v	volume per volume
Mlc	myosin-light chain	w/v	weight per volume
MPD	2-methyl-2,4-pentanediol	w/w	weight per weight
MW	molecular weight	wt	wild-type
MWCO	molecular weight cut-off	X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
Myo	myosin		
NCS	non-crystallographic symmetry	*X. laevis*	*Xenopus laevis*
NMR	nuclear magnetic resonance		

1 Introduction

1.1 mRNA localization in general

An orchestra of gene regulatory events that decides on the fate of a gene or gene product is important to increase the versatility and adaptability of an organism. Such regulatory events occur to a large extent on the transcriptional and translational level. One big advantage of translational regulation is that it allows for precise temporal and spatial control of gene expression and can be achieved for example by asymmetric mRNA distribution [1].

The importance of asymmetric mRNA localization as a general mechanism to target proteins to distinct cellular sites is underlined by the fact that it can be found in many different organisms and different cell types throughout the kingdoms of life. In eukaryotes, specific mRNA localization patterns have been found in animals [2], plants [3] and fungi [4]. Also by far less complex organisms like bacteria were reported to contain localized mRNAs that allow for the uncoupling of transcription and translation [5], [6]. Proper mRNA localization is important to modulate cellular architecture and function and is involved in diverse processes in life such as embryonic patterning and the determination of the body axis [7], [8], cell-fate decisions [9], cell-polarization and migration [4], [10], axonal and dendrite morphogenesis [11] as well as synaptic plasticity [12].

1.1.1 Examples of eukaryotic mRNA localization

The fundamental importance of asymmetric mRNA distribution becomes further clear by the prevalence of localized mRNAs during the embryogenesis of the fruit fly *Drosophila melanogaster (D. melanogaster)*. A transcriptome-wide *in situ* hybridization screen revealed that 71 % of the expressed genes yield subcellularly localized transcripts [13]. Well-known examples are the *oskar*, *nanos* and *bicoid* mRNA that localize to the posterior and the latter to the anterior pole of the oocyte (Figure 1.1 A). Their distribution allows for the generation of morphogen gradients that are the basis for spatial patterning in the developing embryo [14], [15], [7], [16].

A *Drosophila* **embryo**

anterior   posterior

oskar

nanos

bicoid

B *Xenopus* **oocyte**

animal

Vg1

vegetal

C *Saccharomyces cerevisiae*

mother cell

ASH1

daughter cell

D **Fibroblast**

β-actin

Lamellipodium

E **Oligodendrocyte**

myelin lamellae

MBP

Axon

F **Immature Neuron**   **Mature Neuron**

Dendrites   *CamKIIα*

Axon
Myelin Sheath
Growth cone
β-actin   Synapses
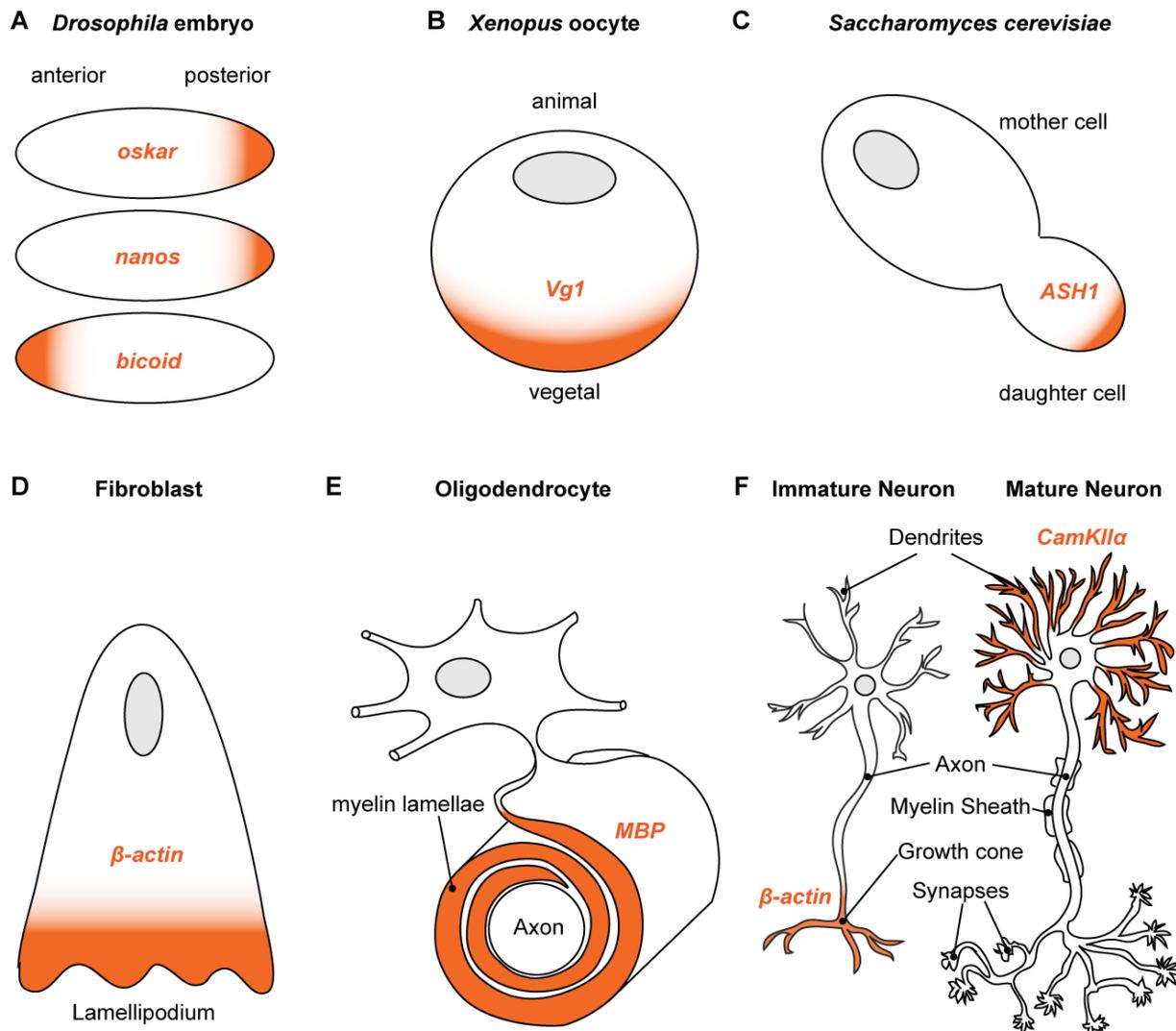
Figure 1.1. Typical examples of localizing mRNAs in different organisms and cell types [1]. Localized mRNAs are depicted in orange. **A)** *oskar* and *nanos* mRNA localize to the posterior and *bicoid* mRNA to the anterior pole of *Drosophila* embryos. **B)** In the clawed frog *X. laevis*, *Vg1* mRNA localizes to the vegetal pole of stage IV oocytes. **C)** In the dividing budding yeast, *ASH1* mRNA localizes to the bud tip of the daughter cell. **D)** *β-actin* mRNA localizes to lamellipodia of mammalian and chick fibroblasts. **E)** *MBP* mRNA localizes to the extensions of the mammalian oligodendrocyte plasma membrane surrounding the axon. Spiraled myelin lamellae form the myelin sheath. **F)** In mammalian immature neurons, *β-actin* mRNA localizes to distal growth cones and in mature, fully developed neurons, *CamKIIα* mRNA localizes to distal dendrites.

Also in the oocytes of the clawed frog *Xenopus laevis* (*X. laevis*) maternal mRNA is asymmetrically accumulated in the cytoplasm [17] (Figure 1.1 B). The vegetally restricted translation of the *Vegetal 1* (*Vg1*) gene, which encodes a transforming growth factor β (TGFβ) signaling molecule, is crucial for endoderm and mesoderm specification during embryogenesis [18].

An asymmetric determination of cell fate can be observed in *Saccharomyces cerevisiae* (*S. cerevisiae*). Here the localization of the *Asymmetric Synthesis of HO 1* (*ASH1*) mRNA to the bud tip of the daughter cell and its localized translation is needed to restrict the mating-type switching exclusively to the mother cell (Figure 1.1 C) [19], [9]. Further details on mRNA localization in budding yeast and especially *ASH1* are described in section 1.3.

In somatic cells like chick fibroblasts, *β-actin* mRNA was initially found to be localized to lamellipodia [20] and later especially assigned to its leading lamellae [21] (Figure 1.1 D). Here, the locally produced β-actin builds the basis for a cytoskeletal network that mediates cell protrusion and motility [22]. Few years ago a genome-wide screen in mouse fibroblasts revealed that upon migratory stimuli at least fifty other transcripts are additionally localized to these pseudopodial protrusions [23].

Another example constitutes the *myelin basic protein* (*MBP*) mRNA. It is localized in the extensions of plasma membranes from mammalian oligodendrocytes and its protein product provides the basis for myelin sheaths that enwrap neuronal axons [24], [25] (Figure 1.1 E).

The last examples shown here derive from mammalian neurons. Beneath its localization in fibroblasts, *β-actin* mRNA is also enriched in growth cones of developing neurons and enables the outgrowth of neurites [26], [27] (Figure 1.1 F). In mature neurons *Calcium/calmodulin-dependent protein kinase II α* (*CamKIIα*) mRNA localizes to distal dendrites [28] where its translation is required for synaptic plasticity and memory consolidation [29] (Figure 1.1 F). Estimated from studies on primary cultures up to hundreds of mRNAs are localized to dendrites [30].

Regarding the implication of mRNA localization in these diverse processes of life it is not surprising that impaired mRNA localization in humans is linked to different pathologies including breast cancer [31] and severe neuronal diseases like mental retardation, epilepsy and autism [32].

1.1.2   Advantages of mRNA localization

Apart from the aforementioned control of gene expression in space and time, there are several other advantages of localizing an mRNA before its actual translation [2], [27], [33]. First, mRNA localization is more efficient than the transport of proteins. Valuable energy is saved since the localization of just one mRNA molecule is sufficient to serve as template for multiple protein copies [2].

Second, the localization of mRNAs prevents proteins from acting ectopically during the transport. Since localized transcripts play important roles for example in embryonic patterning of *D. melanogaster* [8], malfunctions would have disastrous consequences for the whole organism.

Third, proteins can be specifically modulated upon production directly at the site of destination. There, distinct protein properties can be fine-tuned by introducing post-translational modifications. For instance β-actin is arginylated just at the leading edge of fibroblasts, which prevents early filament aggregation and regulates actin function at the site of destination [34].

Forth, the assembly of larger and more intricate protein complexes can be facilitated by local accumulation of mRNAs that code for each of the complex components. This is also the case in the protrusions of fibroblasts for example. Here, all seven mRNAs of the actin-polymerization nucleator (Arp2/3) complex are localized [35]. This results in high local mRNA concentrations, brings translated proteins in close proximity to each other, and increases the likelihood of macromolecular complex formation. Arp2/3 initiates the formation of actin filaments and plays a central role in the branching of actin networks [36].

Lastly, translation can be locally controlled in response to extrinsic signals like neurotransmitters that activate growth cones in synapses or by guidance cues in axonal pathways [37].

1.1.3 Mechanisms of mRNA localization

mRNA localization and the resulting asymmetric distribution of transcripts can be achieved in different ways and even in a combination of those. Although rarely observed, the simplest way to achieve such an asymmetry is local transcription. One example derives from mammalian neuromuscular junctions. Here, the mRNAs for δ- and ε- subunits of the acetylcholine receptor (AChR) are exclusively transcribed in the myonuclei of the synaptic region. While skeletal myofibers have multiple nuclei and form a syncytium, the locally restricted transcription of those mRNAs in nuclei close to synapses leads to receptor production at the place where they are needed [38], [39].

Another way to achieve asymmetry is diffusion-coupled local trapping. This has been shown for the localization of distinct mRNAs in *D. melanogaster* and *X. laevis* oocytes. In *Drosophila* the maternal *nanos* mRNA diffuses from the ovarian nurse cells to the posterior pole of the oocyte where it is actin-dependently anchored to the germ plasm [40]. The resulting gradient of nanos protein has a

morphogenic effect in fly embryos and is essential for abdomen formation. In stage I oocytes of *X. laevis* the germ line RNAs *Xenopus calcium transporter 2* (*Xcat2*) and *Xenopus deleted in azoospermia I* (*XdazI*) are evenly distributed throughout the cytoplasm until they are trapped in the mitochondrial cloud. There they associate with the endoplasmic reticulum (ER) and move with it later on to the vegetal cortex in stage II/III oocytes. In the continuing development they are responsible for the germ plasm inheritance and therefore for germ cell development [41], [42].

The third method to localize mRNAs is the generalized degradation of transcripts in combination with local protection. For instance, this was shown for the heat-shock protein 83 (Hsp83) during the embryogenesis of *D. melanogaster* [43], [44], [45]*.* In the early *Drosophila* embryo maternal *Hsp83* transcripts are initially uniformly spread in the cytoplasm of the oocyte. During development this mRNA gets degraded in the bulk cytoplasm but stays intact in the germ plasm at the posterior pole [43]*.* The key regulator in this system is the Smaug protein. It interacts with smaug response elements (SRE) in the *Hsp83* mRNA [46] and recruits the CCR4-NOT complex that promotes mRNA deadenylation and triggers the 5' decay. Although CCR4-NOT also plays important roles in translational repression, this is not observed in the case of *Hsp83* [47], [48]. Furthermore Smaug and especially its RNA-binding domain were shown to directly interact with the Oskar protein [49]. Since the expression of Oskar is also restricted to the posterior pole, the local protection from degradation most likely derives from the preclusion of *Hsp83*-mRNA binding to Smaug, as it was shown for another maternal mRNA like *nanos* [50].

The best-characterized and probably most abundant mechanism to deliver mRNAs to their destination is active transport driven by molecular motors [27], [51]. It is rapid and can efficiently convey the mRNA cargo over long distances, as it is necessary in axons for instance [52].

1.2   Active mRNA transport

Active mRNA transport is a multi-step process that generally requires *cis*-acting LE-containing mRNAs that guide their transportation [53]. A reoccurring theme is the nuclear priming of mRNAs, which is necessary for their correct cytoplasmic localization [54], [55]. Towards this end trans-acting RBPs associate with LEs in mRNAs and form mRNPs. The binding of trans-acting factors is required for proper cargo localization and translational control [1], [30]. After the initial RNP formation in the nucleus, the complex is reorganized in the cytoplasm and motor proteins are recruited. The latter

actively conduct the cargo delivery along cytoskeletal tracks [52]. At the site of destination anchoring events take place and translational inhibition is released [30], [1].

The composition of such an mRNP can be quite complex, as the examples of mRNPs in neurons show. Here large granules are formed that usually include dozens of proteins [56] and in some cases also hundreds of different mRNAs [57], [58]. This high complexity of motile particles hampers our structural and mechanistic understanding of mRNA localization. However, in *S. cerevisiae* the motor-dependent mRNA transport relies on by far fewer components and is therefore a well-suited model system to study mRNP assembly and function [59].

1.2.1 Diverse *cis*-acting localization elements guide mRNA transport

Distinct regions in mRNAs are used to tag themselves for localization. The existence of such *cis*-acting LEs or so-called zip code elements has been shown many years ago [60]. LEs are recognized by RNA-binding proteins, which connect them to the localization machinery. Several studies showed that a LE fused to a reporter RNA is sufficient to fulfill mRNA localization [61], [62]. Although the majority of LEs reside in 3' untranslated regions (UTRs) of localizing transcripts [53], they are also seldom found in the coding region [61], [53]. Moreover the appearance of a LE can be quite diverse. It ranges from short linear sequences over distinctly folded stem-loop structures to even more intricate entities that adopt complex three dimensional arrangements [53], [63].

A LE in form of a short linear sequence-specific motif resides in the 3' UTR of the *MBP* mRNA that is expressed in oligodendrocytes [64], [65]. Initially it was thought to consist of 21 nucleotides (nt) [66] until it was shown that an 11-nt long stretch called A2 response element (A2RE) is sufficient for RBP binding and mRNA transport [64]. Interestingly, the A2RE element was also found in different other dendritically localized mRNAs including the *CamKIIα* mRNA [67].

The interplay between two short single-stranded sequences is required for *β-actin* mRNA localization in fibroblasts [68], [69]. Here the LE reflects a bipartite motif, which contains a distinct 5' sequence (CGGAC) and a variable 3' element (C/A-CA-C/U) [69]. Patel and colleagues performed NMR experiments and titrated the RNA elements to their RBP zip code-binding protein 1 (ZBP1) to map chemical shifts of amide resonances on the protein. With this technique they deduced an RNA-binding model for the bipartite *cis*-acting element [70] and showed that both sequences have to be appropriately spaced for specific LE recognition [69]. By using this consensus as search motif in

mouse and human 3' UTRs, *spinophilin* mRNA was identified to rely on both ZPB1 and the bipartite LE for dendritic localization [69].

Apart from these short sequence motifs the vast majority of LEs is more complex [53], [63]. Prominent examples are the stem loops in the *ASH1* mRNA from *S.cerevisiae*. Here four LEs act functionally redundant to localize the *ASH1* mRNA [71], [61]. Besides being stem loops their sequences and secondary structures are quite diverse, raising the question what defines them as LEs. However, in budding yeast the core protein components of an mRNP including its RBPs are well-defined [59], thus providing a basis to study specificity-mediating features. A detailed assessment of *ASH1* and its LEs is given in section 1.3.3.

An even more complex structure was described in the 3' UTR of *CamKIIα*, which is localized in neurons. Chemical probing revealed that an intramolecular guanine (G)-quadruplex is necessary and sufficient for neurite localization. This LE was also found in another dendritically localized transcript, the *postsynaptic density protein 95* (*PSD-95*) [72].

Besides the existing linear structures of LEs in literature that are limited to the above-mentioned A2RE [64] and the bipartite *cis*-acting element [68], [69], only two structures of folded stem loops exist. Both derive from *Drosophila* and were elucidated by NMR experiments [73], [74].

The first tertiary structure of an LE responsible for mRNA transport derives from the *fs(1)K10* (*K10*) mRNA in *Drosophila* [73]. This maternal *K10* transcript is transported from the nurse cells to the anterior of the oocyte where it is necessary for the regulation of the dorso-ventral body axis [75], [76]. The LE in *K10* is a 44-nt stem-loop structure termed transport and localization sequence (TLS) [76] and forms an unusual A'-form helix. It harbors two widened major groves with a distinct spatial register that is required for motor-dependent localization [73]. This three-dimensional requirement for signal activity could have never been resolved by biochemical or bioinformatics analysis, which underlines the importance of structure determination techniques for LE analysis.

The second NMR structure comes from an *oskar* LE, which is called spliced *oskar* localization element (SOLE) [74]. As the name implies, splicing at the first exon-exon junction of the *oskar* pre-mRNA is prerequisite for SOLE formation [77] and mRNA localization [78]. The splicing reaction generates a 28-nt stem-loop structure with a widened major grove reminiscent of the *K10* TLS [74]. However to date, just the sequence-independent stem loop of SOLE is assigned as recognition element for the deposition of the exon junction complex, which is also required for active *oskar* transport [74].

1.2.2 RNA-binding proteins recognize LEs and serve as link for other trans-acting factors

The main role of RBPs in mRNP transport is LE recognition and their subsequent linkage to the transport machinery. In this context they also serve as important binding platform for other trans-acting factors like adapter proteins or translational repressors to form a mature mRNP [51]. However, besides LE recognition the same RBP can also act as translational repressor itself [79].

A well-studied RBP involved in active mRNA transport of higher eukaryotes is the predominantly cytoplasmic ZBP1 [70], [80]. Its initial discovery revealed that it is required for the recognition and localization of *β-actin* mRNA in chick embryonic fibroblasts [70]. In the aftermath several orthologues were connected to mRNA localization in different species. These included Vg1 RBP/Vera from *X. laevis* [81], [82], the insulin-like growth factor 2 mRNA-binding protein 1 (IMP1) from human [31] and ZBP1 from rat [83]. ZBP1 contains two N-terminal RNA recognition motifs (RRMs) and four C-terminal heterogeneous nuclear ribonucleoprotein (hnRNP) K homology (KH) domains. While RRM1 and 2 mediate motor association [84], KH domains 3 and 4, are necessary to recognize the bipartite LE of *β-actin* [68], [69]. Since KH3 and 4 reside on opposing sides of the protein, the RNA has to loop around ZBP1 to meet the LE motif requirements [68], [69]. Beneath ZBP1 also the predominantly nuclear RBP ZBP2 is required for proper *β-actin* localization in fibroblasts and neurons [85]. Its initial co-transcriptional binding to the RNA turned out to be prerequisite for efficient ZBP1 association with the nascent transcript [80]. During the course of *β-actin* localization ZBP1 also acts as translational repressor to prevent premature gene expression [79].

The vegetal localization of *Vg1* mRNA in *X. laevis* oocytes relies on the RBPs Vg1 RBP/Vera and hnRNP I that associate with repeated LE sequences in *Vg1* already in the nucleus [55], [86], [87]. Additionally, the hnRNP A/B-D subfamily protein 40LoVe binds the RNA in the nucleus in an hnRNP I-dependent manner [88], [89], [90]. Since 40LoVe alone has low LE specificity, its interaction with *Vg1* is probably supported by at least one associated factor [90]. In the cytoplasm RNP remodeling occurs and further components such as the proline-rich RNA-binding protein (Prrp) and *Xenopus* Staufen (XStau) join the complex [55]. In summary the *Vg1* mRNP harbors a variety of different components, which interact intricately with each other. However, their interplay is still not comprehensively understood.

Another localizing RBP constitutes the fragile-X mental-retardation protein (FMRP). It recognizes the G-quadruplex LE in dendritically localized *CamKIIα* mRNA [72] via an arginine-glycine-rich (RGG) box [91]. Since the deletion of FMRP does not significantly affect transcript localization, additional factors

are required [72]. Experiments in *Drosophila* neurons showed that the adapter protein Bicaudal D (BicD) links FMRP to the localization machinery and regulates motility and function [92].

However, not all RNA-binding proteins do have canonical RNA-binding motifs [93]. One example constitutes Egalitarian (Egl) in *Drosophila* that directly contacts the TLS of *K10* mRNA. Furthermore it is able to interact with the motor co-factor BicD. Since *in vitro* Egl binds LEs not highly specific, the interaction with BicD could mediate specificity *in vivo* [93]. Besides *K10*, Egl and BicD also bind to localizing *Drosophila gurken* mRNA and the pair-rule transcript *hairy*. Other examples of RNA-binding proteins without canonical RNA-binding motifs derive from budding yeast. Here also two proteins are necessary to fulfill specific and synergistic LE-binding. Details on these proteins are described in sections 1.3.4 and 1.3.5.

1.2.3   Motor proteins drive active mRNA transport

Motor proteins of all three motor classes drive active mRNA transport. These classes comprise kinesins, dyneins and myosins that transport their cargoes along cytoskeletal tracks through the cell. They generate force by hydrolyzing adenosine triphosphate (ATP), which results in conformational changes and enables active movement. While kinesin and dynein motors use microtubules to transport their cargo, myosins walk on actin filaments. The direction of their movement is basically determined by their subclass. While most kinesins and myosins drive plus-end directed transport [94], [95], dyneins mostly fulfill minus-end directed transport [95] although they are also capable of bidirectional movement [96]. One common requirement for motor molecules that transport mRNAs is their dimerization via a coiled-coil domain. Apart from dedicated mRNAs that are transported with this motor protein repertoire also a variety of other cargoes like macromolecular protein complexes, vesicles, viruses, centrosomes and even whole organelles are known to be carried [95], [27], [94], [52].

Kinesin-1-mediated mRNA transport is observed in neurons for example [56]. Here a highly polarized microtubule network with minus ends starting in the cell body and plus ends ranging to the periphery ensures proper anterograde transcript localization [52]. Evidence derives from isolated kinesin-associated granules that localized to dendrites and contained amongst different proteins and mRNAs also *CamKIIα* mRNA, FMRP, Purine-rich element-binding protein A (Pur-alpha) and Staufen [56]. Another kinesin-dependency was demonstrated for *MBP*-mRNA localization in oligodendrocytes [97].

The interplay between two different motor species working in a directional cue was shown for the localization of the *Vg1* mRNA in *Xenopus* oocytes. Here, the initial unidirectional transport to the vegetal cortex is dependent on dynein. It is coupled to a second non-directional step, which is mediated by kinesin-1. Since microtubules in the vegetal pole possess opposing polarities, kinesin motors can act there to refine cargo distribution [98].

Also the localization of *oskar* mRNA in *Drosophila* depends on different motors that act sequentially to finally deposit the transcript at the posterior. Recent findings showed that *oskar* has a second LE with similarities to the TLS of *K10* mRNA [99]. This oocyte entry signal (OES) directs *oskar* minus-end transport from the nurse cells to the oocyte in a dynein-dependent manner [99]. Since Egl and BicD are indispensable for *oskar* localization [100] and both proteins interact with the similar *K10* LS [93], they were also suggested to mediate the dynein contact in case of OES [99]. Once in the oocyte SOLE takes over and guides *oskar* with the help of kinesin motors to the posterior pole [77]. Directionality in this process relies on a bias in microtubule orientation [101]. Zimyanin and colleagues showed in living oocytes that *oskar* mRNPs are generally transported by kinesin in all directions on a weakly polarized microtubule network. Since a small majority of microtubule plus ends point towards the posterior pole, the cargo transport is slightly biased and *oskar* can accumulate at its site of destination [101].

Budding yeast mRNA transport relies on the type V myosin motor Myo4p. It actively transports its *ASH1*-cargo RNP along actin filaments from the minus ends in the mother cell to the plus ends in the daughter cell. Although one experimental set-up showed that the recruitment of more motors resulted in increased transport efficiency [102], processivity-mediating features are still discussed. Further details on Myo4p are described in section 1.3.6.

More recently also a kinesin-5 microtubule motor KIF11 was shown to directly interact with ZBP1 in mouse embryonic fibroblasts. Knocking down KIF11 or deleting its interaction site to the RBP resulted in impaired *β-actin* mRNA localization to the leading edge of fibroblasts [84]. Microtubule-association was also shown for the localization of *β-actin* and ZBP1 to growth cones in neurons [103]. In live-cell imaging experiments ZBP granules moved in a bidirectional fashion, which indicated that both kinesin and dynein motors participate in transport [103]. Interestingly, later on it was shown that the type V myosin MyoVa is able to regulate the transport dynamics of ZBP1, probably by influencing the availability of ZBP1 for active transport on microtubules [104].

Apart from these mechanisms also microtubule-associated proteins can regulate transport. One prominent example is the protein Tau [105]. It decorates microtubules in neurons and stabilizes them [106]. Tau was reported to tune the velocity of kinesins, but also to inhibit kinesin-mediated transport in an isoform-dependent manner [107], [108]. For dynein motors there is also evidence that they reverse their direction upon encountering Tau [107], [109].

1.3   mRNA localization in *S. cerevisiae*

In the past 20 years *S. cerevisiae* turned out to be a well-suited, simple model system to study the basic principles of mRNA localization, mRNP assembly and function [59]. Apart from several hundred mRNAs that are associated with organelle membranes of mitochondria or the ER [110], currently more than 30 mRNAs are known to be actively localized to specific subcellular regions in budding yeast [111], [112], [113], [114]. These include for instance different bud-localized transcripts like the *Increased Sodium Tolerance 2* (*IST2*) mRNA that encodes for Ist2p, which tethers the ER to the plasma membrane [115] or the *cell Wall integrity and Stress response Component 2* (*WSC2*) mRNA, where the encoding protein is involved in the maintenance of cell wall integrity and the recovery from heat shock [116], [117]. Another example is the *Endosomal Adaptor of Rsp5p 1* (*EAR1*) mRNA. Its protein product Ear1p assists in tagging membrane proteins for their destiny in the vacuole [118]. In contrast to several localized transcripts with unknown functions [113], to date the best-studied example however is the bud-tip targeted *ASH1* mRNA [59].

1.3.1   Mating-type switching

Proper *ASH1*-mRNA transport is required to inhibit mating-type switching in homothallic *S. cerevisiae* strains [9], [19]. In haploid budding yeast cells the *MAT* locus determines the cell's mating type, which consists either of the *MAT*a or the *MAT*α allele [119]. Mating-type switching from *MAT*a to *MAT*α, or the other way round, exclusively occurs within the mother cell. Here the *homothallic switching* (*HO*) gene is expressed during late G1 phase and results in the production of the HO endonuclease [120]. This enzyme generates a site-specific double-strand break at the *MAT* locus that is prerequisite for homologous recombination with adjacent genetic donor cassettes carrying complete copies of the silenced mating-type genes [121], [122], [123]. Thereby one *MAT* allele is replaced by a DNA sequence encoding for the opposite allele, resulting in intra-chromosomal gene

conversion and therefore mating-type switching. In the daughter cell this process is inhibited by the expression of the HO specific repressor gene *ASH1*. Its respective protein product is called "asymmetric synthesis of HO 1" (Ash1p) [124], [125].

1.3.2 The *SHE* genes

At the same time when Ash1p was found to be responsible for the determination of the daughter cell mating type, five *SWI5-dependent HO expression* (*SHE*) genes were discovered. Their protein products were shown to be necessary for the asymmetric distribution of Ash1p [126]. Shortly afterwards two studies revealed that it is the *ASH1* mRNA and not Ash1p that is actively transported to the bud tip [9], [19]. The involved genes in this process (*SHE1-SHE5*) encode cytoplasmic proteins that were just partially known before. While at that time She2p, She3p and She4p were completely uncharacterized, it was known that *SHE1* encodes the type V myosin motor protein Myo4p [127], and *SHE5* is translated into the formin Bud neck involved 1 (Bni1p) [128]. The latter promotes the assembly of actin filaments to actin cables and is therefore required for the establishment of a stable mother-bud axis [129]. In the following years She4p/Dim1p was shown to interact with Myo4p and to be essential for its motor function [130], [131]. Further experiments ascribed She2p to be an RNA-binding protein [132] and already suggested a central role in building an active mRNP-transport complex together with the Myo4p-associated She3p protein. The latter acts as adapter molecule to deliver *ASH1* to its site of destination [111], [132], [133]. Details on each of the components that form the SHE complex (*ASH1* mRNA, She2p, She3p and Myo4p) are described in the following sections.

1.3.3 *ASH1* mRNA – Cargo with localization elements

Together with the discovery of *ASH1*-mRNA transport also its first *cis*-acting localization signal was assigned. Initial fluorescence *in situ* hybridization experiments showed that the 3' UTR of *ASH1* is sufficient to localize a chimeric reporter RNA to the daughter cell. Therefore a *cis*-acting element was deduced to reside in the 3' UTR of *ASH1* that guides mRNA localization [9]. Nevertheless it was also suspected that *ASH1* harbors additional *cis*-acting regions outside the 3' UTR since its replacement with a 3' UTR of an unrelated transcript just slightly reduced *ASH1*-mRNA localization [9]. First moving particles containing an MS2-aptamer insertion in front of the *ASH1*-3' UTR were observed in

living yeast by monitoring the bound GFP-tagged MS2-coat protein [134]. In the following year two concurrent studies revealed that in total four secondary structural elements forming stem-loop structures, termed E1, E2A, E2B and E3, act functionally redundant to localize the *ASH1* mRNA [71], [61]. While E3 is the only LE residing in the 3' UTR, the others are spread over the coding region (Figure 1.2). Each element alone was sufficient to localize a reporter mRNA to the bud [61].

Apart from their central role in guiding the *ASH1* transport, E1 and E3 are involved in mediating the translational repression during the transport from the mother to the bud. In this respect the KH-Domain 1 protein (Khd1p) binds to CNN repeats in the E1 LE [135], [136] and the Pumilio-homology domain Family 6 protein (Puf6p) associates with PUF consensus sequences in E3 [137]. Both trans-acting proteins operate as translational repressor during cytoplasmic mRNA transport and ensure that Ash1p is not ectopically expressed [136], [137], [138].

The fact that the same She2p-She3p localization complex recognizes different mRNAs and even different LEs in one mRNA like *ASH1* [139] urges to find common features amongst them. Such findings should help to unravel how an LE is specifically recognized for mRNA transport. Although being frequently addressed in the past, essential common denominators are still not well understood. It was generally shown that the stem-loop structures of the *ASH1* LEs are required for proper mRNA localization [140]. However, a direct comparison of the LEs showed only minimal similarities in sequence or secondary structure features [141] (Figure 1.2). One attempt to define a She2p-binding consensus motif utilized a secondary structure search algorithm and tested 22 bud-localized mRNA sequences [142]. In this study Olivier and colleagues postulated a motif containing a combination of a single-stranded CGA base triplet and a conserved cytosine in an opposing loop. Although they assigned this motif in all four *ASH1* LEs it was just found in two more localizing mRNAs *IST2* and *YMR171c* and is lacking for instance in the *EAR1* LE (Figure 1.2). On closer examination the motif can also occur in an inversed 3' to 5' orientation, with sequence permutations in the base triplet or a varying location of the bipartite motif in the LE secondary structure (Figure 1.2). Another study found a single-stranded CG dinucleotide in a short sequence motif of a stem loop to be important for She2p-She3p recognition. Since other nucleotides also contributed to the recognition in a context-dependent manner, variations in the sequence or structure are tolerated in some cases [143].

Although the last-mentioned motif was just found in a subset of LEs in *ASH1*, it lead also to the discovery of two LEs in the bud-localized *WSC2* mRNA [143]. Nevertheless several other localizing

mRNAs still await the definition of their LEs and a likely common recognition feature. Therefore the previously defined motifs are not sufficient to cover the variety of LEs that are recognized by the She2p-She3p complex. Future structural work on the transport complex will be needed to complement bioinformatics and interaction studies to elucidate how an LE is specifically recognized for mRNA transport.
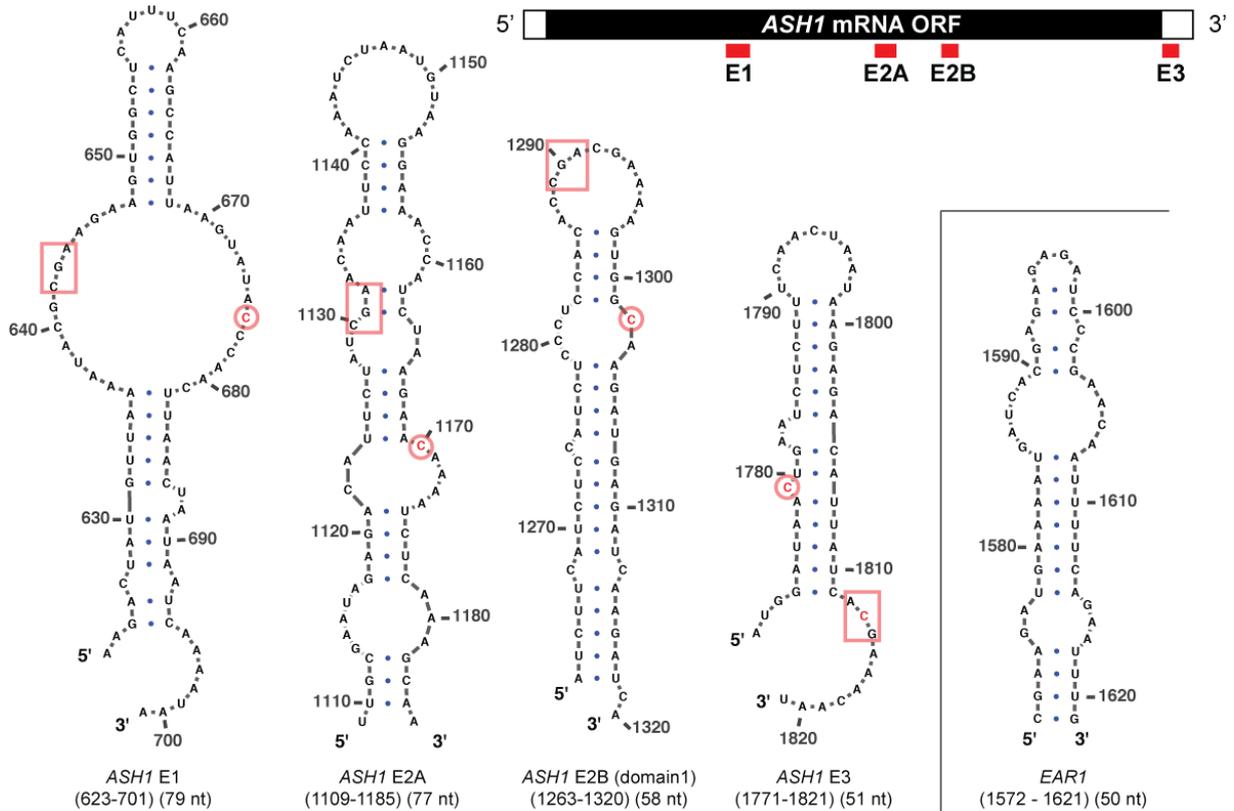
Figure 1.2 Secondary structure predictions of the *ASH1* and *EAR1* localization elements. The upper right part shows a schematic representation of the *ASH1* mRNA and the position of its localization elements E1, E2A, E2B and E3. While the three first-mentioned LEs reside in the open-reading frame (ORF) the latter element is located directly after the stop-codon in the 3' UTR. The lower part of the picture shows secondary structure predictions [144] of LEs from *ASH1* and *EAR1* that are bound by the She2p-She3p complex. The stems contain internal loops and mismatches resulting in bulges. One of the postulated recognition motifs is highlighted. It comprises a CGA base triplet (red boxes) and a cytosine (red circles) at a certain distance [142]. Of note, in E3 the whole motif has an inversed 5' to 3' orientation and the base triplet a permutation in its sequence. The *EAR1* LE completely lacks this motif.

1.3.4   She2p – RNA-binding protein and central component of the *ASH1* mRNP

The RNA-binding protein She2p [132],[133] has a central role in the processes involved in *ASH1*-mRNP formation and its localization [59], [145]. It is able to shuttle with the RNA from its place of birth [146] in the nucleus via the nucleolus to the cytoplasm [147], [148]. In this process the

nucle(ol)ar transition is essential for proper translational repression of the *ASH1* transcript [148], [149]. In the cytoplasm She2p constitutes one of the core components of the *ASH1* mRNP due to its interaction with LE-containing RNAs and She3p [132].

The crystal structure of She2p was solved in 2004 and shows two globular monomers (A and B) that form a 2-fold symmetric homodimer. Each 28-kDa monomer is composed of five α-helices arranged in a bundle with an additional protruding helix at the side, which is also called helix E (hE) [150] (Figure 1.3 A).

Although not having a canonical RNA-binding domain a positively charged surface region comprising the "basic helical hairpin" was mapped on She2p to be responsible for RNA-binding. Mutations of single amino acids in this area impaired or abolished LE binding *in vitro* and mRNA localization *in vivo* [150], [151] (Figure 1.3 B).

Subsequently, analytical ultracentrifugation and SAXS experiments revealed that She2p forms an elongated tetramer with a head-to-head conformation in solution [152] (Figure 1.3 C). This tetramer is required for RNA binding, mRNP assembly and mRNA localization [102], [152]. Further static-light scattering (SLS) experiments clarified that one She2p tetramer is able to bind two LE RNAs [145]. Although not resolved in the crystal structure, also the very C-terminus of She2p (residues 241-246) is involved in RNA binding and consequently its deletion results in loss of mRNA localization *in vivo* [139].

While She2p alone binds LEs with low affinity and specificity, it is also able to directly interact RNA-independently with She3p *in vitro* [139] (Figure 1.3 C, D). However, just the combination of all three components together forms a highly stable synergistic ternary complex with increased affinity and specificity for localizing elements [139]. Pull-down experiments with She2p mutants (F195A, L196A) or (Q197A, E198A, I199A) and full-length She3p indicated that these regions, residing in a flexible loop between the protruding helix and the globular central part of She2p, are necessary for the She2p-She3p interaction [139] (Figure 1.3 C, D). In a recently solved crystal structure [153] eight additional She2p residues located in two of the core helices (Figure 1.3 D) interacted with a short She3p peptide (see also Figure 1.4 B). Also the protruding helices in She2p resemble functionally important parts since their deletion (residues 174-183) abolishes the She2p-She3p interaction, reduces RNA binding and abrogates synergistic RNA binding *in vitro* as well as *ASH1* localization *in vivo* [139] (Figure 1.3 C, D).
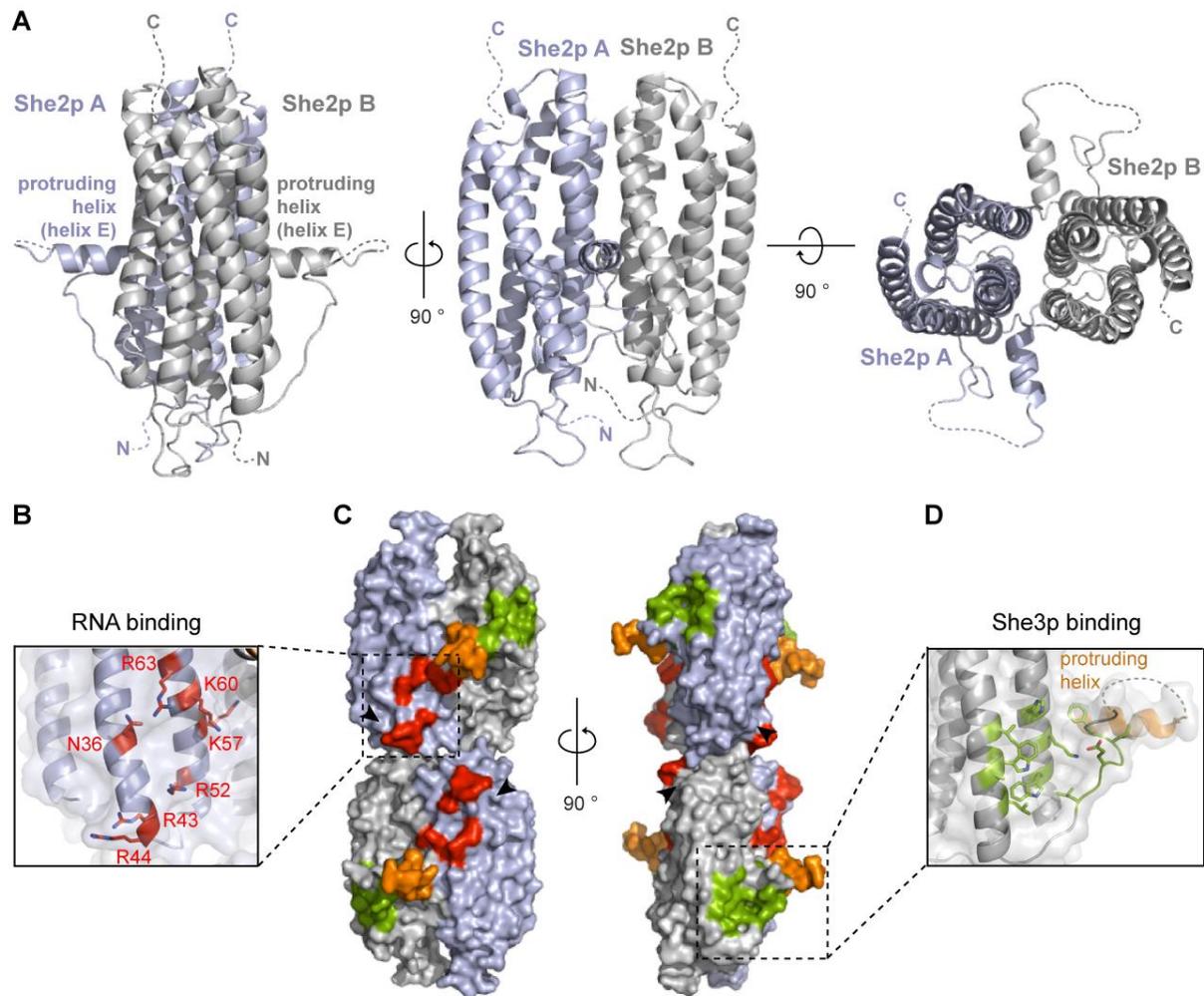
Figure 1.3: Structure of She2p and a subset of important RNA and She3p interaction sites. A) Cartoon representations of the She2p homodimer crystal structure (PDB-ID: 1XLY) [150]. The homodimer is built of the more complete monomer from the crystal structure. Monomers (She2p A and B) are colored in light-blue and gray. She2p is shown in side view (left), front view (middle) and top view (right). Dashed lines indicate regions that were not visible in the structure. **B)** Close-up of the dashed box in C) shows the basic helical hairpin [150] in which amino acids that are involved in RNA binding are highlighted in red [151], [150]. **C)** Surface representation of the elongated tetramer [152] in front (left) and side view (right) with red surfaces deriving RNA-binding amino acids mentioned in B). The protruding helices are colored in orange and the She3p-interacting residues are colored in green [153], [139]. Black arrowheads mark positions of C-terminal tails that were absent in the crystal structure. Dashed boxes are magnified in B) and D). **D)** Close-up of the She3p binding region comprising residues F195-L200, H150, K153, W157, L211, A214, W215 and I218, depicted in green [153], [139]. Amino acids 174-183 are colored in orange and belong to helix E.

1.3.5 She3p – Adaptor to the transport machinery and RNA-binding protein

The strictly cytoplasmic She3p [139] is essential for *ASH1*-mRNA localization since it acts as an adaptor protein that links She2p to the Myo4p motor in the transport complex [111], [132], [133]. The secondary structure of the 425 amino acid-containing She3p is predicted to be α-helical in its N-terminal half, whereas its C-terminal half is predicted to be mainly unstructured (Figure 1.4 A, Appendix Figure 5.1).

The N-terminal half of She3p interacts with the C-terminal tail of Myo4p, thereby forming a constitutive cytoplasmic co-complex [132], [154], [155], [156]. Recently a combination of chemical cross-linking, isothermal titration calorimetry (ITC) and SLS experiments revealed that She3p dimerizes via its N-terminus and forms a stable heterotrimeric complex with a Myo4p monomer *in vitro* [145]. This finding was confirmed by two crystal structures just one year later [157] (Figure 1.4 B). They showed that the N-terminal α-helical half of She3p forms a pseudo coiled-coil, in which the hydrophobic region comprising residues 149-193 is bound by the globular tail of Myo4p (Figure 1.4 A, B).

The unstructured C-terminal half of She3p was implicated in She2p binding [132], [133]. Pull-down and surface plasmon resonance experiments verified their direct association [139]. This finding was supported by mutational studies on She2p, as described in the previous section 1.3.4, and an additional She3p (L364A, V367A) mutant that eliminated the She2p-She3p interaction [139]. Recently, the crystal structure of a co-complex consisting of She2p and an eight amino acid fragment of She3p was solved [153]. It demonstrates that a conserved LPGV-hook motif (residues 364-368) in She3p binds to a conserved hydrophobic pocket in She2p (Figure 1.4 A, C). As evident from pull-down experiments with mutant She3p proteins, each motif residue is essential for the She2p-She3p interaction [153]. These structural insights and earlier *in vitro* reconstitution experiments prove that four She3p molecules bind four She2p molecules [145], [153].

Apart from being an adapter molecule between Myo4p and She2p, She3p was also shown to be an RNA-binding protein [139]. Comparable to She2p, She3p alone binds LE RNAs with low affinity and specificity [139]. However, in the ternary complex both proteins directly bind to RNA and mediate synergistic, high-affinity RNA binding. UV cross-linking experiments with this ternary complex in combination with mass-spectrometric analysis revealed that RNA binding of She3p is attributed to a 7-nt long peptide (residues 334-340) [139] (Figure 1.4 A). Further experiments assessed two C-terminal She3p fragments (354-425) and (334-425) regarding their ability to form a ternary complex.

Just the latter fragment was able to fully support synergistic RNA binding together with She2p [139] (Figure 1.4 A).

Interestingly She3p (334-425) contains three residues (343, 348 and 361) that were identified as phosphorylation sites. These residues were assessed by the two phosphor-mimicking mutants She3p S348E and She3p (S343E, S361E). Although both retained the ability to bind She2p and Myo4p, they were defective for *ASH1*-mRNA localization. This indicates a negative regulation mechanism of She3p upon phosphorylation [158].

In summary the cytoplasmic protein She3p serves as motor-adapter molecule and is responsible for synergistic LE-RNA recognition with She2p. Thereby She3p stably incorporates localizing mRNAs into active transport particles.
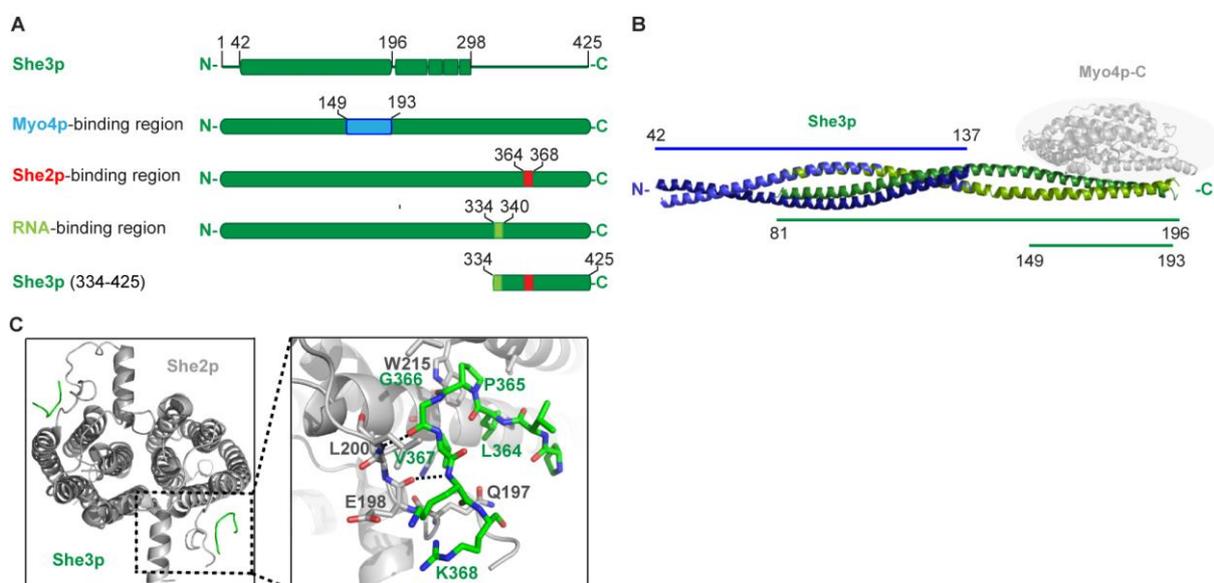
Figure 1.4: Adapter protein She3p and its SHE complex-interaction sites. A) Schematic representations of She3p (dark green). Upper illustration shows the full-length She3p (1-425) secondary structure prediction according to **Appendix Figure 5.1**. Region 42-298 is mainly predicted to be α-helical, whereas most of the C-terminal half of the protein (298-425) is predicted to be unstructured. The Myo4p-binding region (149-193) [157] is depicted in blue, the She2p-binding region (364-368) is colored in red and RNA-binding region (334-340) is shown in lime green [139]. She3p (334-425) is the shortest She3p fragment known to fully support synergistic ternary complex formation with She2p **B)** Superposition of She3p (42-137) (dark and light blue) and the She3p (81-196, dark and lime green):Myo4p-C (1016-1471, light gray) co-complex crystal structures (PDB-ID: 4LL7, 4LL8, respectively). The N-terminal half of She3p forms a pseudo coiled-coil. The C-terminal globular domain of Myo4p binds to She3p comprising residues 149-193 [157]. **C)** Left panel: Top view of the She2p dimer (gray) bound by two She3p (364-368) peptides (green) (PDB-ID: 4WNL). Right panel: Close-up of boxed area from left panel, rotated -30 ° around the y-axes. The She3p ^{364}LPGV^{367}-hook interacts with She2p residues W215, L200 and E198 amongst others [153].

1.3.6 Myo4p – The transporting motor protein

The type-V myosin motor Myo4p is the executing transport agent for the directional *ASH1*-mRNA localization [159]. It is monomeric in isolation [155], [156], [160] and its structural organization comprises an N-terminal motor domain, a central neck-region and a globular C-terminal tail [161], [162].

While the motor domain enables the molecule to attach to actin and is able to bind ATP [161] and She4p [130], [131], the neck region builds the binding platform for calmodulin (CaM) and myosin light chains (Mlc) [161]. The C-terminal tail of type V myosins is dedicated to the binding of cargo molecules like the adaptor protein She3p that links the motor to *ASH1* and She2p [111], [132], [145], [157], [163]. In line with this notion the deletion of the tail domain leads to impaired *ASH1*-mRNA localization *in vivo* [163].

Monomeric Myo4p is non-processive [160] and also its constitutively bound She3p dimer does not induce motor oligomerization [145]. However, it was shown that active movement requires at least two, most likely dimeric, motor proteins [102], [160], [164]. A combination of different experiments clarified that the *ASH1*-bound She2p tetramer associates with two She3p-Myo4p heterotrimers, which in turn assembles the SHE-core complex and induces Myo4p activation [145]. Apart from these findings, the relevance of cargo RNA for complex assembly and motility is still controversially discussed [165], [159], [164]. While Sladewski and colleagues showed that the RNA is required to form a stable and processive transport complex [159], a study from Krementsova and co-workers suggested that RNA is not necessary for motor activation [164].

1.3.7 Summarizing model for *ASH1*-mRNA localization in *S. cerevisiae*

The process of *ASH1*-mRNA localization from the mother cell to the bud tip of the daughter starts with the She2p protein in the cytoplasm. Its non-classical nuclear localization signal triggers the protein import through importin alpha Srp1p into the nucleus [149] (Figure 1.5 – step 1).

Here, She2p binds co-transcriptionally to the nascent *ASH1* mRNA by interacting with the elongating RNA polymerase II machinery via its transcription elongation-factor complex-components Spt4p and Spt5p [146] (Figure 1.5 – step 2).

In vitro binding studies revealed that the early nuclear pre-mRNP consisting of She2p and the *ASH1* mRNA has only low specificity and stability [152], [139]. At first glance this appeared inconsistent

with its *in vivo* function since She2p selectively delivers LE-containing mRNAs from their site of transcription to their destination at the bud tip of the daughter cell. However, this seeming contradiction was recently resolved by showing that the strictly nucle(ol)ar protein LOCalization of mRNA 1 (Loc1p) [166], [167] binds with high affinity to both She2p and *ASH1* in the pre-mRNP. In this ternary complex Loc1p stabilizes their otherwise transient interaction and increases the specificity for LE-containing mRNAs by roughly five-fold [168]. Before leaving the nucleus the pre-mRNP undergoes a nucleolar transition and is loaded with additional translational repressors [148], [149] (Figure 1.5 – step 3).
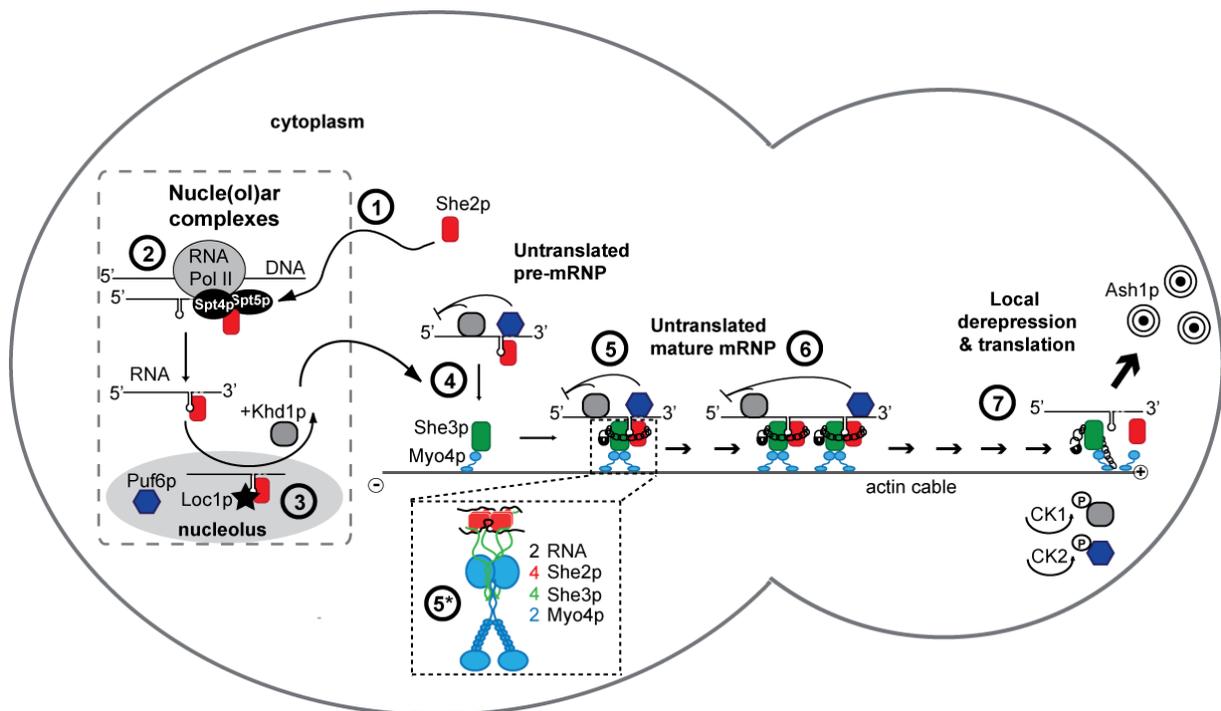
Figure 1.5: Model of *ASH1*-mRNA localization in budding yeast. Step 1: She2p (red) enters the nucleus. **Step 2:** She2p binds to the elongating RNA polymerase II machinery by interacting with the transcription elongation factors Spt4p and Spt5p (black ovals). It binds to stem-loop forming LE structures in the *ASH1* mRNA. **Step 3:** The pre-mature mRNP formed of She2p, *ASH1*, and Loc1p (black star) translocates through the nucleolus, where it is thought to be loaded additionally with the translational repressor Puf6p (dark blue). Also Khd1p joins the complex in the nucleus. **Step 4:** After nuclear export the translationally repressed pre-mRNP associates with the Myo4p-bound She3p protein (depicted in light blue and green, respectively). **Step 5:** A synergistic, high-affinity complex that stably binds *ASH1* LEs is formed. Joining of She3p induces Myo4p dimerization, yielding a mature mRNP that is capable of active cargo transport along the actin cable to its plus end. **Magnification 5*:** The stoichiometry of the SHE complex consisting of *ASH1* mRNA, She2p, She3p and Myo4p is 2:4:4:2. **Step 6:** Multiple LEs in *ASH1* mRNA give rise to a larger mRNPs. **Step 7:** After the transport to the bud tip of the daughter cell local derepression occurs. CK1 and CK2 phosphorylate Khd1p and Puf6p, respectively, and the complex is thought to disassemble. Locally restricted translation and Ash1p production takes place, which inhibits mating type switching in the daughter cell.

Loc1p, which also represses translation [166], is assisted by Puf6p that directly binds to the *ASH1* mRNA [137], [139], [168]. The latter interacts with the eukaryotic translation initiation factor 5B (eIF5B, or Fun12p in budding yeast) and consequently interferes with the joining of the ribosomal subunits [138]. Additionally, Khd1p binds with one of its KH-domains to *ASH1* in the pre-mRNP and interacts with eIF4G1, which hinders cap-dependent translation initiation [136], [135], [148]. However in contrast to Puf6p and Loc1p, Khd1p does not enter the nucleolus [167], [148].

During nuclear export mRNP reorganization occurs. Loc1p is replaced at the nuclear pore by the strictly cytoplasmic Myo4p-associated She3p protein [168] (Figure 1.5 – step 4).

In the cytoplasm She3p assumes the role of Loc1p. Its synergistic RNA binding together with She2p yields a highly stable complex with further increased affinity and specificity for LE-containing mRNAs [139], [168]. Upon joining of She3p its bound and otherwise monomeric Myo4p [160], [156] dimerizes [155], [164]. This leads to motor activation and the formation of a mature mRNP that walks along the actin cable to its plus end at the tip of the daughter cell [164] (Figure 1.5 – step 5).

In the mature transport complex one She2p tetramer binds two LE-containing RNAs, two dimers of She3p and two Myo4p motors [145] (Figure 1.5 – Magnification 5*). Moreover due to the fact that *ASH1* has more than just one LE, the SHE core complex is able to multimerize into even larger particles [145] (Figure 1.5 – step 6).

When reaching the bud tip of the daughter cell, Khd1p and Puf6p are phosphorylated by the membrane associated casein kinase 1 (CK1) and CK2, respectively. This reduces the affinity for the *ASH1* mRNA and results in translational activation [169], [138] (Figure 1.5 – step 7).

Taken together, mRNP maturation is necessary in *S. cerevisiae* to ensure that just LE-containing transcripts are transported as translationally silenced particles. Since no complex structure of such a transport particle is available to date also a structural understanding of how specificity for LE-containing RNAs is mediated is still missing.

1.4 Objectives of the thesis

To date, the assembly- and processivity-mediating features of motile particles for asymmetric mRNA localization are not well understood. At the beginning of my thesis, studies on mRNA-transport in yeast comprised experiments with RNP complexes purified from cell extracts [102] or with partial complexes lacking the RNA cargo [164]. More recent *in vitro* reconstitution experiments clarified the stoichiometric ratios within a SHE complex [145] and provided a basis to study mRNP assembly and motility under well-defined conditions. The first part of my study contributed to the assessment of SHE complex assembly and motility. Which features activate motility? Is it the RNA-binding protein She2p or its cargo RNA *ASH1*? Besides *ASH1*-mRNA localization, She2p is involved in the transport of about 30 other transcripts into the daughter cell [112], [143]. This fact raises the question whether different RNAs can be transported in one SHE complex particle. Furthermore it was shown that an RNA with two localization elements induces clustering of SHE core complexes *in vitro* [145]. Is such a complex multimerization also seen for actively transported particles, and if so, what effect would it have on its processivity? The experimental set-up, molecular cloning, protein purification, RNA production and single-particle motility assays, the latter performed by a collaboration partner, are described in section 3.1.

An assessment of the stability of the Myo4p-containing transport complex has not been performed yet. This gap of knowledge was a matter of discussion in literature and lead to speculations. In a commentary, Sladewski and colleagues suggested that protein-protein interactions in the transport complex, without RNA and at low ionic strength would lead to artificial particle movement [165]. Furthermore they speculated that a complex of Myo4p, She2p and She3p under physiological salt concentrations would not be stable *in vitro* without its cargo RNA. To eliminate these uncertainties, the second part of my study directly addresses the question if the presence of LE RNA would be required for complex assembly at physiological salt concentrations and temperature. What effect on stability do varying salt concentrations and temperatures have? Pull-down and SLS experiments, as well as analytical size-exclusion chromatography (SEC) were used to answer these questions and are shown in section 3.2.

The third and most extensive part comprises structural studies on the SHE core complex consisting of She2p, She3p and the *ASH1* mRNA. Since at the beginning of this thesis only the X-ray structure of She2p alone [150] was available, I aimed to co-crystallize the complex of She2p and She3p together with RNA, as well as sub-complexes of it. Concerning high-resolution structures of localization

elements the only known examples of structured stem loops were the *fs(1)K10* mRNA, and SOLE in the *oskar* mRNA of *D. melanogaster*, which were elucidated by NMR experiments [73], [74]. How would the first crystal structure of such a localization element in yeast look like? Are the structures of LEs in *ASH1*, *EAR1* or *WSC2* comparable? Only comparing their structural properties could shed light on their operating modes since they do not share sequence or secondary structure similarity but are all specifically recognized by the same She2p-She3p protein complex. To obtain insights into the assembly path of a transport particle I also aimed to solve the crystal structure of the nuclear co-complex of She2p and an RNA LE in absence of the cytoplasmic She3p.

How is specific LE recognition by the mature cytoplasmic transport machinery achieved? Referring to this question a recently published structure of She2p in complex with a short She3p peptide [153] did not give any answer. It is known that cargo specificity in the ternary, cytoplasmic complex consisting of She2p, She3p and the *ASH1* mRNA is increased compared to the nuclear co-complex consisting of She2p and the RNA [139]. Nevertheless structural determinants that mediate this increase of specificity remain poorly understood. In this context a crystal structure of the ternary complex would yield the missing information. Structural methods like SAXS, NMR and X-ray crystallography, complemented with biochemical assays like electrophoretic mobility-shift assays (EMSAs) were used to accomplish these tasks as described in section 3.3.

2 Materials and Methods

2.1 Chemicals and consumables

Commonly used chemicals were purchased from Sigma Aldrich (Deisenhofen, Germany), Roth (Karlsruhe, Germany), Merck (Darmstadt, Germany) or AppliChem (Darmstadt, Germany) unless stated otherwise. Ingredients to prepare bacterial growth media derived from Becton Dickinson (Heidelberg, Germany). Nucleotides, enzymes and molecular weight markers for molecular biology techniques were ordered at New England BioLabs (Frankfurt, Germany), Agilent Technologies (Waldbronn, Germany) or Thermo Scientific (Bonn, Germany). Cell culture flasks and plates were provided by Greiner Bio-One (Frickenhausen, Germany) and Thermo Scientific (Bonn, Germany). Membrane filter devices derived from Merck Millipore (Darmstadt, Germany) or GE Healthcare (Munich, Germany). The latter also provided liquid chromatography material, devices and columns with the exception of amylose resin, which was ordered at New England BioLabs (Frankfurt, Germany). N-terminally acetylated, C-terminally amidylated and HPLC-purified peptides were purchased from JPT (Berlin, Germany). Unless stated otherwise, chemicals for crystallization were ordered in the highest available purity at the above-mentioned companies. Hampton Research (Aliso Viejo, USA) and Qiagen (Hilden, Germany) provided commercially available crystallization screens and tools.

2.2 Organisms and strains

Strain	Description/Genotype	Source
Bacteria		
E. coli XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacI^q ZΔM15* Tn*10* (Tet^r)]	Agilent Technologies
E. coli Stellar^{TM}	F$^-$ *endA1 supE44 thi-1 recA1 relA1 gyrA96 phoA* Φ80*d lacZΔ* M15, *Δ(lacZYA-argF)* U169, *Δ(mrr*-hsdRMS-*mcr*BC) *ΔmcrA* λ$^-$	Clontech
E. coli BL21 Star (DE3)	F$^-$ *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm rne131* (DE3)	Stratagene
E. coli BL21 Gold (DE3) pRARE	B F$^-$ *dcm$^+$* Hte *ompT hsdS*(r_B^- m_B^-) *gal endA* Hte [*arg*U *ile*X *leu*W *pro*L *met*T *gly*T *tyr*U *thr*U (Cam^r)]	Novagen

Strain	Description/Genotype	Source
E. coli B834 (DE3)	F⁻ *ompT hsdS*_B(r_B⁻ m_B⁻) *gal dcm met* (DE3)	Novagen
E.coli M15 DZ291	F⁻ Φ80ΔlacM15 thi lac⁻ mtl⁻ recA⁺ (Kan^R)	Qiagen
E.coli DH10Bac	F⁻*mcr*A Δ(*mrr*-hsdRMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara,leu*)7697 *gal*U *gal*K λ⁻ *rps*L *nup*G/bMON14272/pMON7124	Invitrogen
Insect cells		
*Sf*21	Derived from *Spodoptera frugiperda* (*Sf*) ovarian cells. Used to isolate and propagate recombinant baculovirus stocks.	Invitrogen
High five	Derived from the *Trichopulsia ni* cell line. Used to express recombinant proteins upon baculovirus infection.	Invitrogen

Table 2.1: Strains and organisms.

2.3 Plasmids

Vector	Application/Description	Tag	Resistance	Source/Reference
Bac10:KO₁₆₂₉	recombination substrate for insect cell transfection	-	Cam, Kan	Reference [170]
pBSM	RNA expression in bacteria	tRNA-Met scaffold	Amp	Reference [171]
pET28a	protein expression in bacteria	N-terminal His₆	Amp	Merck Millipore
pETM-41	protein expression in bacteria	N-terminal MBP-TEV[†]	Kan	Dr. Arie Geerlof, PEPF, HMGU
pETM-43	protein expression in bacteria	N-terminal MBP-3C[#]	Kan	Dr. Arie Geerlof, PEPF, HMGU
pFastBac^{TM} Dual	protein expression in insect cells	-	Amp, Gent	Invitrogen
pFBDM	protein expression in insect cells	-	Amp, Gent	Reference [172]
pGEX-4T-1	protein expression in bacteria	N-terminal-GST-Thr*	Amp	Pharmacia
pGEX-6P-1	protein expression in bacteria	N-terminal-GST-3C[#]	Amp	GE Healthcare
pOPIN-J	protein expression in bacteria/ insect cells	N-terminal His₆-GST-3C[#]	Amp	Addgene
pOPIN-M	protein expression in bacteria/ insect cells	N-terminal His₆-MBP-3C[#]	Amp	Addgene
pOPIN-S3C	protein expression in bacteria/ insect cells	N-terminal His₆-SUMO-3C[#]	Amp	Addgene
pQE	protein expression in bacteria	N-terminal His₆	Amp	Qiagen
pRS405	yeast integrative vector	-	Amp	Reference [173]
YCplac22	yeast shuttle vector	-	Amp	Reference [174]
YCplac111	yeast shuttle vector	-	Amp	Reference [174]

Table 2.2: Base Vectors. † TEV stands for tobacco etch virus (TEV)-protease cleavage site, [#] 3C indicates human rhino virus (HRV) 3C-protease cleavage site. * Thr resembles thrombin cleavage site.

Plasmid name	Description	PCR primers	Source/ Cloning strategy
	Plasmids deriving from different sources		
AHP7	pGEX-6P-1-*MYO4(978-1471)*	-	Reference [155]
p01	pGEX-*SHE2*	-	Reference [150]
p02	pET28a_*eGFP_SHE2*	-	Reference [145]
p03	pGEX-6P-1_*SHE2(6-246 C-S)*	-	Dr. Marisa Müller
p04	pGEX-6P-1_*SHE2(6-243 C-S)*	-	Dr. Marisa Müller
p05	pGEX-6P-1_*SHE2(6-240 C-S)*	-	Dr. Marisa Müller
p10	pGEX-6P-1_*SHE2_N36S_R63K*	-	Reference [148]
pGEX/3C	pGEX_*3C protease*	-	Dr. Arie Geerlof, PEPF, HMGU
pRJ350	YCplac22_*SHE3_myc3*	-	Prof. Jansen, Universität Tübingen
pRJ1978	YCplac111_*SHE2_E172A_F176A*	-	Prof. Jansen, Universität Tübingen
RHP27	pFastBacDual-*SHE2-SHE3His6*	-	Reference [139]
RHP82	pGEX-6P1_*SHE2(6-246)*	-	Dr. Roland Heym
RHP85	pGEX-*SHE3(334-425)His6*	-	Reference [139]
RHP94	pGEX-6P-1-*SHE3(343-425)His6*	-	Dr. Roland Heym
RHP143	pETM-43-*SHE3(334-425)His6*	-	Dr. Roland Heym
RHP145	pBSMrna-*2x-ASH1-E3-77*	-	Reference [145]
RHP149	pBSMrna-*IST2*	-	Reference [145]
RHP163	pET28a-*CMD1*	-	Reference [145]
RHP166	pRS-405-*ASH1-3'*	-	Reference [145]
RHP167	pET28a-*MLC1*	-	Reference [145]
RHP169	pFBDM-*FLAG-MYO4-CMD1-MLC1-SHE3-HIS6-SHE4*	-	Reference [145]
	Plasmids generated in this study		
FE#2	pETM-43_*SHE3(334-375)*	O176/P5	Restriction sites: NcoI/EcoRI
FE#3	pETM-43_*SHE3(334-375)His6*	O176/P2	Restriction sites: NcoI/EcoRI
FE#4	pGEX-6P-1_*SHE3(334-375)*	P4/P5	Restriction sites: BamHI/EcoRI
FE#5	pGEX-6P-1_*SHE3(334-375)His6*	P4/P2	Restriction sites: BamHI/EcoRI
FE#6	pETM-43_*SHE3(334-405)*	O176/P6	Restriction sites: NcoI/EcoRI
FE#7	pETM-43_*SHE3(334-405)His6*	O176/P3	Restriction sites: NcoI/EcoRI
FE#8	pGEX-6P-1_*SHE3(334-405)*	P4/P6	Restriction sites: BamHI/EcoRI
FE#9	pGEX-6P-1_*SHE3(334-405)His6*	P4/P3	Restriction sites: BamHI/EcoRI
FE#10	pGEX-6P-1_*SHE3(334-382)*	P4/P7	Restriction sites: BamHI/EcoRI
FE#11	pGEX-6P-1_*SHE3(334-382)His6*	P4/P8	Restriction sites: BamHI/EcoRI
FE#12	pGEX-6P-1_*SHE3(334-389)*	P4/P9	Restriction sites: BamHI/EcoRI
FE#13	pGEX-6P-1_*SHE3(334-389)His6*	P4/P10	Restriction sites: BamHI/EcoRI
FE#14	pGEX-6P-1_*SHE3(334-398)*	P4/P11	Restriction sites: BamHI/EcoRI
FE#15	pGEX-6P-1_*SHE3(334-398)His6*	P4/P12	Restriction sites: BamHI/EcoRI
FE#16	pETM-43_*SHE3(334-382)His6*	O176/P8	Restriction sites: NcoI/EcoRI
FE#17	pETM-43_*SHE3(334-389)His6*	O176/P10	Restriction sites: NcoI/EcoRI
FE#18	pETM-43_*SHE3(334-398)His6*	O176/P12	Restriction sites: NcoI/EcoRI
FE#19	pETM-41_*SHE3(334-405)*	-	Restriction sites: NcoI/EcoRI, cut and paste from FE#6
FE#20	pET28a_*eGFP_SHE2_N36S*	P21/P22	QuickChange on p02
FE#21	pET28a_*eGFP_SHE2_R63K*	P23/P24	QuickChange on p02
FE#22	pET28a_*eGFP_SHE2_N36S_R63K*	P23/P24	QuickChange on FE#20
FE#23	pET28a_*eGFP_SHE2ΔhE*	P25/P26	QuickChange on p02
FE#24	pFBDual_*SHE3-His6_SHE2_N36S_R63K*	P29/P30	Restriction sites: SmaI/NcoI
FE#25	pFBDual_*SHE3-His6_SHE2ΔhE* (deletion of aa174-183)	P29/P30	Restriction sites: SmaI/NcoI
FE#26	pGEX-6P-1_*SHE3(334-Δ(344-359)405)His6*	P43/P44	QuickChange on FE#9
FE#27	pETM-43_*SHE3(334-Δ(344-359)405)*	P43/P44	QuickChange on FE#6

Plasmid name	Description	PCR primers	Source/ Cloning strategy
FE#28	pGEX-6P-1_*SHE3(334-Δ(345-359)405)His6*	P41/P42	QuickChange on FE#9
FE#29	pETM-43_*SHE3(334-Δ(345-359)405)*	P41/P42	QuickChange on FE#6
FE#30	pOPIN-J_*SHE2(6-246)C-S_GGSGG_SHE3(331-375)*	P48/P49, P51/P53	InFusion: 3 point PCR on p03 and FE#6
FE#31	pOPIN-J_*SHE2(6-246)C-S_(GGSGG)x2_SHE3(331-375)*	P48/P50, P52/P53	InFusion: 3 point PCR on p03 and FE#6
FE#32	pOPIN-J_*SHE2(6-246 C-S)_GGSGG_SHE3(331-405)*	P48/P49, P51/P54	InFusion: 3 point PCR on p03 and FE#6
FE#33	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)*	P48/P50, P52/P54	InFusion: 3 point PCR on p03 and FE#6
FE#34	pOPIN-J_*SHE2(6-246 C-S)_GGSGG_SHE3(331-Δ(345-359)405)*	P48/P49, P51/P54	InFusion: 3 point PCR on p03 and FE#28
FE#35	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-Δ(345-359)405)*	P48/P50, P52/P54	InFusion: 3 point PCR on p03 and FE#28
FE#36	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)*	P48/P55, P56	InFusion: 2x primer extension PCR on p03
FE#37	pOPIN-M_*SHE2(6-246 C-S)_GGSGG_She3p(331-375)*	P48/P49, P51/P53	InFusion: 3 point PCR on p03 and FE#6
FE#38	pOPIN-M_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-375)*	P48/P50, P52/P53	InFusion: 3 point PCR on p03 and FE#6
FE#39	pOPIN-M_*SHE2(6-246 C-S)_GGSGG_She3p(331-405)*	P48/P49, P51/P54	InFusion: 3 point PCR on p03 and FE#6
FE#40	pOPIN-M_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)*	P48/P50, P52/P54	InFusion: 3 point PCR on p03 and FE#6
FE#41	pOPIN-M_*SHE2(6-246 C-S)_GGSGG_SHE3(331-Δ(345-359)405)*	P48/P49, P51/P54	InFusion: 3 point PCR on p03 and FE#28
FE#42	pOPIN-M_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-Δ(345-359)405)*	P48/P50, P52/P54	InFusion: 3 point PCR on p03 and FE#28
FE#43	pOPIN-M_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)*	P48/P55, P56	InFusion: 2x primer extension PCR on p03
FE#44	pOPIN-M_*SHE2(6-246 C-S)_GGSGG_SHE3(331-343)*	P48/P55, P56	InFusion: 2x primer extension PCR on p03
FE#45	pOPIN-S3C_*SHE2(6-246 C-S)_GGSGG_SHE3(331-375)*	P48/P49, P51/P53	InFusion: 3 point PCR on p03 and FE#6
FE#46	pOPIN-S3C_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-375)*	P48/P50, P52/P53	InFusion: 3 point PCR on p03 and FE#6
FE#47	pOPIN-S3C_ *SHE2(6-246 C-S)_GGSGG_SHE3(331-405)*	P48/P49, P51/P54	InFusion: 3 point PCR on p03 and FE#6
FE#48	pOPIN-S3C_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)*	P48/P50, P52/P54	InFusion: 3 point PCR on p03 and FE#6
FE#49	pOPIN-S3C_*SHE2(6-246 C-S)_GGSGG_SHE3(331-Δ(345-359)405)*	P48/P49, P51/P54	InFusion: 3 point PCR on p03 and FE#28
FE#50	pOPIN-S3C_*SHE2(6-246 C-S)_ (GGSGG)x2_SHE3(331-Δ(345-359)405)*	P48/P50, P52/P54	InFusion: 3 point PCR on p03 and FE#28
FE#51	pOPIN-S3C_*SHE2(6-246 C-S)_GGSGG_SHE3(331-343)*	P48/P55, P56	InFusion: 2x primer extension PCR on p03
FE#52	pOPIN-S3C_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)*	P48/P55, P56	InFusion: 2x primer extension PCR on p03
FE#53	pETM-43_*SHE3(334-405)Δ(352-354)*	P65/P66	QuickChange on FE#6
FE#54	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2*	P48/P73	InFusion: Primer extension PCR on p03
FE#55	pOPINJ_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(382-405)*	P48/P50, P64/P54	InFusion: 3 point PCR on p03 and FE#6
FE#56	pOPINJ_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(342-374)Δ(352-354)*	P48/P50, P67/68	InFusion: 3 point PCR on p03 and FE#53
FE#57	pOPINJ_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(362-369)*	P48/P50, P69	InFusion: 2x primer extension PCR

Plasmid name	Description	PCR primers	Source/ Cloning strategy
			on p03
FE#58	pGEX-4T-1_*SHE3(342-374)Δ(352-354)*	P70/P71	Restriction sites: BamHI/EcoRI
FE#59	pGEX-4T-1_*SHE3(382-405)*	P72/P6	Restriction sites: BamHI/EcoRI
FE#60	pOPIN-S3C_*SHE3(342-374)Δ(352-354)*	P74/P68	InFusion: Template: FE#53
FE#61	pOPIN-S3C_*SHE3(382-405)*	P75/P54	InFusion: Template: FE#6
FE#62	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_N333M*	P48/P91, P90/P54	InFusion: 3 point PCR on FE#33
FE#63	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_K340M*	P48/P93, P92/P54	InFusion: 3 point PCR on FE#33
FE#64	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_S343M*	P48/P95, P94/P54	InFusion: 3 point PCR on FE#33
FE#65	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)_K340A*	P48/P101	InFusion: Template: FE#36
FE#66	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)_R342A*	P48/P103	InFusion: Template: FE#36
FE#67	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)_R341A*	P48/P102	InFusion: Template: FE#36
FE#68	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_N339A*	P48/P83, P82/P54	InFusion: 3 point PCR on FE#33
FE#69	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_N338A*	P48/P85, P84/P54	InFusion: 3 point PCR on FE#33
FE#70	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_N335A*	P48/P87, P86/P54	InFusion: 3 point PCR on FE#33
FE#71	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_N333A*	P48/P89, P88/P54	InFusion: 3 point PCR on FE#33
FE#72	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)_N333A*	P48/P97	InFusion: Template: FE#36
FE#73	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)_N335A*	P48/P98	InFusion: Template: FE#36
FE#74	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)_N338A*	P48/P99	InFusion: Template: FE#36
FE#75	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)_N339A*	P48/P100	InFusion: Template: FE#36
FE#76	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-343)_K340A,R341A,R342A*	P48/P107	InFusion: Template: FE#65
FE#77	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_K340A*	P48/P109, P108/P54	InFusion: 3 point PCR on FE#33
FE#78	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_R341A*	P48/P111, P110/P54	InFusion: 3 point PCR on FE#33
FE#79	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_R342A*	P48/P113, P112/P54	InFusion: 3 point PCR on FE#33
FE#80	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)_K340A,R341A,R342A*	P48/P114, P115/P54	InFusion: 3 point PCR on FE#33
FE#81	pOPIN-J_*SHE2(6-246 C-S)_ (GGSGG)x3_GGS_SHE3(344-405)*	P48/P116, P117/P54	InFusion: 3 point PCR on FE#33
FE#82	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x4_GGS_SHE3(344-405)*	P48/P119, P118/P54	InFusion: 3 point PCR on FE#81
FE#83	pOPIN-J_*SHE3(331-405)_(364-368)A*	P122/P121, P120/P54	InFusion: 3 point PCR on FE#33
FE#84	pOPIN-J_*SHE3(331-405)_(340-342)A_(364-368)A*	P122/P121, P120/P54	InFusion: 3 point PCR on FE#33 and FE#80
FE#85	pOPIN-J_*SHE3(331-405)*	P122/P54	InFusion: Template: FE#33
FE#86	pOPIN-J_*SHE3(331-405)_(340-342)A*	P122/P54	InFusion: Template: FE#80
FE#87	pOPIN-S3C_*SHE3(331-405)_(364-368)A*	P122/P121, P120/P54	3 point PCR on FE#33
FE#88	pOPIN-S3C_*SHE3(331-405)_(340-342)A_(364-368)A*	P122/P121, P120/P54	InFusion: 3 point PCR on FE#33 and FE#80
FE#89	pOPIN-S3C_*SHE3(331-405)*	P122/P54	InFusion: Template: FE#33
FE#90	pOPIN-S3C_*SHE3(331-405)_(340-342)A*	P122/P54	InFusion: Template: FE#80
FE#91	pOPIN-S3C_*SHE3(331-343)*	P122/P56	InFusion: Template: FE#33
FE#92	pOPIN-S3C_*SHE3(342-374)*	P74/P68	InFusion: Template: FE#33

Plasmid name	Description	PCR primers	Source/ Cloning strategy
FE#93	pOPIN-S3C_*SHE3(344-381)*	P123/P125	InFusion: Template: FE#33
FE#94	pOPIN-S3C_*His6-SUMO_L*	P126/P127	InFusion: 1:1 primer as insert
FE#95	pGEX-6P-1_*SHE2_E172A_F176A*	P128/P129	QuickChange on p01
FE#96	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-346)*	P48/P136	InFusion: Template: FE#33
FE#97	pOPIN-J_*SHE2(6-246 C-S)_(GGSGG)x2_SHE3(331-405)(K340A_R341A_Y345A)*	P48/P137, P138/P54	InFusion: 3 point PCR on FE#33
FE#98	pOPIN-S3C_*SHE3(331-405)(K340A_R341A_Y345A)*	P122/P54	InFusion: Template: FE#97
FE#99	pOPIN-S3C_*SHE3(331-425)His6*	P122/P139	InFusion: Template: RH84
FE#100	pOPIN-S3C_*SHE3(331-425)*	P122/P140	InFusion: Template: RH84
FE#101	Ycplac22_*SHE3(340-346)A_myc3*	P144/P146, P147/P145	3 point PCR on pRJ350, EcoRI/SalI
FE#102	Ycplac22_*SHE3(364-368)A_myc3*	P144/P148, P149/P145	3 point PCR on pRJ350, EcoRI/SalI
FE#103	Ycplac22_*SHE3(340-346)A(364-368)A_myc3*	P144/P148, P149/P145	3 point PCR on FE#101, EcoRI/SalI
FE#104	pGEX-6P-1_*SHE2_E172A_F176A_StuI*	P152/P153	QuickChange on FE#95
FE#105	pOPIN-J_*SHE3(331-405)K340A*	P122/P54	InFusion: Template: FE#77
FE#106	pOPIN-J_*SHE3(331-405)R341A*	P122/P54	InFusion: Template: FE#78
FE#107	pOPIN-J_*SHE3(331-405)R342A*	P122/P54	InFusion: Template: FE#79
FE#108	pOPIN-S3C_*SHE3(331-405)K340A*	P122/P54	InFusion: Template: FE#77
FE#109	pOPIN-S3C_*SHE3(331-405)R341A*	P122/P54	InFusion: Template: FE#78
FE#110	pOPIN-S3C_*SHE3(331-405)R342A*	P122/P54	InFusion: Template: FE#79
FE#111	pETM-43_*SHE3(334-405)Y345A*	P156/P157	QuickChange on FE#6
FE#112	pOPIN-J_*SHE3(331-405)Y345A*	P122/P54	InFusion: Template: FE#11
FE#113	pOPIN-S3C_*SHE3(331-405)Y345A*	P122/P54	InFusion: Template: FE#111
FE#114	pETM-43_*SHE3(334-405)F344A*	P159/P160	QuickChange on FE#6
FE#115	pOPIN-J_*SHE3(331-405)F344A*	P122/P54	InFusion: Template: FE#114
FE#116	pOPIN-S3C_*SHE3(331-405)F344A*	P122/P54	InFusion: Template: FE#114
FE#117	Ycplac22_*SHE3(K340A, R341A)_myc3*	P144/P164, P163/P145	3 point PCR on pRJ350, EcoRI/SalI
FE#118	Ycplac22_*SHE3(L364A, V367A)_myc3*	P144/P166, P165/P145	3 point PCR on pRJ350, EcoRI/SalI
FE#119	Ycplac22_*SHE3(K340A, R341A, Y345A)_myc3*	P167/P168	QuickChange on FE#117

Table 2.3: Plasmids.

2.4 Oligonucleotides

Primer name	Sequence (5'→3')	Primer name	Sequence (5'→3')
DNA oligonucleotides for molecular cloning			
O176[#]	AAACCATGGGTAATAGCAGTAACAATAAAAG	**P92**	GGTAATAGCAGTAACAATATGAGAAGAAGTTTCTATAC
P2	TTTTGAATTCCTAGTGATGGTGATGGTGATGTTTTCTTACTGAGGCAGTT	**P93**	CGCAGTATAGAAACTTCTTCTCATATTGTTACTGCTATT
P3	TTTTGAATTCCTAGTGATGGTGATGGTGATGTCTTGTAACACCGGGAGGC	**P94**	GCAGTAACAATAAAAGAAGAATGTTCTATACTGCGTCG
P4	AAAAGGATCCGGTAATAGCAGTAACAATA	**P95**	GCAAAGGCGACGCAGTATAGAACATTCTTCTTTTATTGTTAC
P5	TTTTGAATTCCTATTTTCTTACTGAGGCAGTTC	**P97**	ATGGTCTAGAAAGCTTTAACTTCTTCTTTTATTGTTACTGCTATTACCAGCGGAGGATCCACCACT
P6	TTTTGAATTCCTATCTTGTAACACCGGGAGGCACTG	**P98**	ATGGTCTAGAAAGCTTTAACTTCTTCTTTTATTGTTACTGCTAGCACCATTGGAGGATCC
P7	TTTTGAATTCCTATTTACTACTGCTCGAGCTTGGTTTTC	**P99**	ATGGTCTAGAAAGCTTTAACTTCTTCTTTTATTGGCACTGCTATTACCATT

Primer name	Sequence (5'→3')	Primer name	Sequence (5'→3')
P8	TTTTGAATTCCTAGTGATGGTGATGGTGATGTTTACTACTGCTCGAGCTTGG	P100	ATGGTCTAGAAAGCTTTAACTTCTTCTTTTAGCGTTACTGCTATTACC
P9	TTTTGAATTCCTAGTTATTATGCGTTACATTTGTTTTACTACTG	P101	ATGGTCTAGAAAGCTTTAACTTCTTCTTGCATTGTTACTGCTAT
P10	TTTTGAATTCCTAGTGATGGTGATGGTGATGGTTATTATGCGTTACATTTGTTTTAC	P102	ATGGTCTAGAAAGCTTTAACTTCTTGCTTTATTGTTACTGCT
P11	TTTTGAATTCCTATGAGATTGTGGGAGATGTACTTGG	P103	ATGGTCTAGAAAGCTTTAACTTGCTCTTTTATTGTTACT
P12	TTTTGAATTCCTAGTGATGGTGATGGTGATGTGAGATTGTGGGAGATGTACTTGG	P107	ATGGTCTAGAAAGCTTTAACTTGCTGCTGCATTGTTACTGCTAT
P21	CATCTTATATTCACGTGCTGAGCAAGTTCATCAGTCATTTGCG	P108	GGTAATAGCAGTAACAATGCAAGAAGAAGTTTCTATAC
P22	CGCAAATGACTGATGAACTTGCTCAGCACGTGAATATAAGATG	P109	CGCAGTATAGAAACTTCTTCTTGCATTGTTACTGCTATT
P23	CTACTTTGATTAAATTTGTTAAGAAATTGAAATTTTACAACGATTGTGTGTTAAGCTA	P110	TAATAGCAGTAACAATAAAGCTAGAAGTTTCTATACTGCG
P24	TAGCTTAACACACAATCGTTGTAAAATTTCAATTTCTTAACAAATTTAATCAAAGTAG	P111	CGCAGTATAGAAACTTCTAGCTTTATTGTTACTGCTATTA
P25	TCTCCGATATTCGTTCCGTCAACTTCTAAATCTAGCAGGTTGCTGC	P112	TAGCAGTAACAATAAAAGAGCAAGTTTCTATACTGCGTCG
P26	TGCTAGATTTAGAAGTTGACGGAACGAATATCGGAGAAACTG	P113	CGACGCAGTATAGAAACTTGCTCTTTTATTGTTACTGCTA
P29	AAAACCATGGCTCGAGTCAGTTTTTCAATTTACCAAATTTGTCATGC	P114	CGACGCAGTATAGAAACTTGCAGCTGCATTGTTACTGCTATTACC
P30	TTTTCCCGGGATGAGCAAAGACAAAGATATCAAAGTCACTCC	P115	GGTAATAGCAGTAACAATGCAGCTGCAAGTTTCTATACTGCGTCG
P35	TAATACGACTCACTATAGGGCGAATTGGGTACC	P116	GCTGCCACCACCGCCGCTGCCACCTCCACCACTTCCACC
P36	CTGCGCAGGAGAAGTTATTAGAATGATTTCACTTC	P117	AGCGGCGGTGGTGGCAGCGGCGGTGGTGGCAGTTTCTATACTGCGTCGCCTT
P41	AACAATAAAAGAAGAAGTTTCGCATCCCCGGTTTTGCCTGGTG	P118	GGGGGAAGTGGGGGAGGTGGCAGTTTCTATACTG
P42	ACACCAGGCAAAACCGGGGATGCGAAACTTCTTCTTTTATTG	P119	TCCCCCACTTCCCCCACCGCCGCTGCCACCAC
P43	CAGTAACAATAAAAGAAGAAGTGCATCCCCGGTTTTGCCTGGTG	P120	AATCTGCATCCCCGGTTGCGGCTGCTGCTGCAAGAACTGCCTCAGTAAG
P44	ACACCAGGCAAAACCGGGGATGCACTTCTTCTTTTATTGTTAC	P121	CTTACTGAGGCAGTTCTTGCAGCAGCAGCCGCAACCGGGGATGCAGATT
P48	AAGTTCTGTTTCAGGGCCCGGATATCAAAGTCACTCCTGGAACTTCTG	P122	AAGTTCTGTTTCAGGGCCCGTCCTCCAATGGTAATAGCAG
P49	TCCACCACTTCCACCGTTTTTCAATTTACCAAATTTGTCGTGCCA	P123	AAGTTCTGTTTCAGGGCCCGTTCTATACTGCGTCGCC
P50	TCCACCACTTCCACCTCCACCACTTCCACCGTTTTTCAATTTACCAAATTTGTCGTGCCA	P125	ATGGTCTAGAAAGCTTTAACTACTGCTCGAGCTTGG
P51	GGTGGAAGTGGTGGATCCTCCAATGGTAATAGCAGTAACAATAAAAGAAGAAGTTTC	P126	AAGTTCTGTTTCAGGGCCCGTAAAGCTTTCTAGACCAT
P52	GGTGGAAGTGGTGGAGGTGGAAGTGGTGGATCCTCCAATGGTAATAGCAGTAACAATAAAAGAAGAAGTTTC	P127	ATGGTCTAGAAAGCTTTACGGGCCCTGAAACAGAACTT
P53	ATGGTCTAGAAAGCTTTATTTTCTTACTGAGGCAGTTCTTTAACACC	P128	AACCTGCTAGATTTAGCAGTTGTTCAAGCTGCCATTAAATGCGCAGATG
P54	ATGGTCTAGAAAGCTTTATCTTGTAACACCGGGAGGCAC	P129	CTGCGCATTTAATGGCAGCTTGAACAACTGCTAAATCTAGCAGGTT
P55	ATTGTTACTGCTATTACCATTGGAGGATCCACCACTTCCACCTCC	P136	ATGGTCTAGAAAGCTTTAAGTATAGAAACTTCTTCTTTTATTG
P56	ATGGTCTAGAAAGCTTTAACTTCTTCTTTTATTGTTACTGCTATTACCATTGGAGGA	P137	TAGCGAAACTTCTTGCTGCATTGTTACTGCTATTACC
P64	GGTGGAAGTGGTGGAGGTGGAAGTGGTGGAAAAACAAATGTAACGCATAATAACGATCCAAGTAC	P138	TGCAGCAAGAAGTTTCGCTACTGCGTCGCCTTTGCTG
P65	TACTGCGTCGCCTTTGCTGTCTATTCCAAAATCTGC	P139	ATGGTCTAGAAAGCTTTAGTGATGGTGATGGTGATGGGATTGGGC
P66	GGGATGCAGATTTTGGAATAGACAGCAAAGGCGACGC	P140	ATGGTCTAGAAAGCTTTAGGATTGGGCCCCGTGAACAAC
P67	GGTGGAAGTGGTGGAGGTGGAAGTGGTGGAAGAAGTTTCTATACTGCGTCGCCTTTGC	P144	CGACGGCCAGTGAATTC
P68	ATGGTCTAGAAAGCTTTATCTTACTGAGGCAGTTCTTTTAACACCAGG	P145	CCTGCAGGTCGACTCTA

Primer name	Sequence (5'→3')	Primer name	Sequence (5'→3')
P69	ATGGTCTAGAAAGCTTTATCTTTTAACACCAGGCAAAACCGGTCCACCACTTCCACCTCCACCACTTCCACC	P146	TGCCGCTGCCGCTGCCGCTGCATTGTTACTGCTATTACCATT
P70	AAAAGGATCCAGAAGTTTCTATACTGCGTCG	P147	CAATGCAGCGGCAGCGGCAGCGGCAGCGTCGCCTTTGCTGTCATCG
P71	TTTTGAATTCCTATCTTACTGAGGCAGTTC	P148	GTTCTTGCAGCAGCAGCCGCAACCGGGGATGCAGATTTTG
P72	AAAAGGATCCAAAACAAATGTAACGCATAATAACG	P149	TTGCGGCTGCTGCTGCAAGAACTGCCTCAGTAAG
P73	ATGGTCTAGAAAGCTTTATCCACCACTTCCACCTCCACCACTTCCACCGTTTTTCAATTTACCCAAATTTG	P152	TTTGGTAAATTGAAAAACTAAGGCCTCCCGGGTCGACTCGAGCGG
P74	AAGTTCTGTTTCAGGGCCCGAGAAGTTTCTATACTGCGTCGCCTTTGC	P153	GCGGCCGCTCGAGTCGACCCGGGAGGCCTTAGTTTTTCAATTTACC
P75	AAGTTCTGTTTCAGGGCCCGAAAACAAATGTAACGCATAATAACGATCCAAGTAC	P156	AAAAGAAGAAGTTTCGCTACTGCGTCGCCTT
P82	CCAATGGTAATAGCAGTAACGCTAAAAGAAGAAGTTAA	P157	GCAAAGGCGACGCAGTAGCGAAACTTCTTCTT
P83	GCTTTAACTTCTTCTTTTAGCGTTACTGCTATTACC	P159	CAATAAAAGAAGAAGTGCCTATACTGCGTCGC
P84	CCTCCAATGGTAATAGCAGTGCCAATAAAAGAAGAAGT	P160	GCAAAGGCGACGCAGTATAGGCACTTCTTCTTTTA
P85	GCTTTAACTTCTTCTTTTATTGGCACTGCTATTACCATTG	P163	GTAATAGCAGTAACAATGCAGCAAGAAGTTTCGCTACT
P86	GGTGGATCCTCCAATGGTGCTAGCAGTAACAATAAAAG	P164	GCAAAGGCGACGCAGTAGCGAAACTTCTTGCTGCATT
P87	CTTCTTTTATTGTTACTGCTAGCACCATTGGAGGATCC	P165	GTTGCGCCTGGTGCTAAAAGAACTGCCTCAGTAAG
P88	GTGGAAGTGGTGGATCCTCCGCTGGTAATAGCAGTAACA	P166	TTTAGCACCAGGCGCAACCGGGGATGCAGATT
P89	CTTTTATTGTTACTGCTATTACCAGCGGAGGATCCACCACTTC	P167	AGTAACAATGCAGCAAGAAGTTTCGCTACTGCGTCGCC
P90	GGAAGTGGTGGATCCTCCATGGGTAATAGCAGTAAC	P168	CAGCAAAGGCGACGCAGTAGCGAAACTTCTTGCTGCAT
P91	GTTACTGCTATTACCCATGGAGGATCCACCACT		

	DNA oligonucleotides as templates for *in vitro* transcription		
P19	AATTTAATACGACTCACTATAGG	P59	AGTTGAAGTTTTCGTGTCCAATTTCGTGGACGTTTTTGAACTAACACTGCATAGACTTTCATCCTTAGCTATAGTGAGTCGTATTAAATT
P20	ATTGTTTCGTGATAATGTCTCTTATTAGTTGAAAGAGATTCAGTTATCCATCCCTATAGTGAGTCGTATTAAATT	P60	CAAATTCTGAAAATTGTTCGGGATCTCTCTCGTGATCATTTTCATCTTCGGCATAGACTTTCATCCTTAGCTATAGTGAGTCGTATTAAATT
P31	AATTTAATACGACTCACTATAGGGATGCTGGAGTGAAGAATGTC	P61	GAGATTTTGTTCTTAGATGGTTTGCATAGACTTTCATCCTTAGCAATTGTTCGATAGAATGTCTCTATAGTGAGTCGTATTAAATT
P32	TTTTGGTTTTTCCTTTGCCGCCGATGAAGAGGAAGAGGATTC	P62	ATCTTGCCACTTTTCGTCGGTGTGGAGGGAGATGCATAGCACTTTCATGCCTTAGCTATAGTGAGTCGTATTAAATT
P33	AATTTAATACGACTCACTATAGGGAGAGAATTGATACATGGATAACTG	P63	CATCTTGCCACTTTTCGTCGGTGTGGAGGGAGATGGCATAGACTTTCATCCTTAGCTATAGTGAGTCGTATTAAATT
P34	GTAGTTTATTTAGCACAGACAAGGAGAGAAATGTACAATTGTTTCGTG	P96	CGTGATAATGTCTTTCGATTCAGTTATCTATAGTGAGTCGTATTAAATT
P37	AATTTAATACGACTCACTATAGGGCTACGTAGCTCAGTTG	P130	ATTGTTTCGTGATAATGTCTCTTATTAGTTGAAAGAGATTCTGTTATCCATCTATAGTGAGTCGTATTAAATT
P38	GTAGTTTATTTAGCACAGACAAGGTGGTGGCTACGACGGGATTC	P131	ATTGTTTCGTGATAATGTCTCTTATTAGTTGAAAGAGATTCGGTTATCCATCTATAGTGAGTCGTATTAAATT
P45	AATTTAATACGACTCACTATAG	P132	ATTGTTTCGTGATAATGTCTCTTATTAGTTGAAAGAGATTCAGTTATCCATCTATAGTGAGTCGTATTAAATT
P46	ATCTTGCCACTTTTCGTCGGTGTGGAGGGAGATGCATAGACTTTCATCCTTAGCTATAGTGAGTCGTATTAAATT	P133	ATTGTTTCCTGATAATGTCTCTTATTAGTTGAAAGAGATTCAGTTATCCATCTATAGTGAGTCGTATTAAATT
P47	CCACTTTTCGTCGGTGTGGGCATAGACTTTCATCCTTAGCTATAGTGAGTCGTATTAAATT	P134	ATTGTTTCATGATAATGTCTCTTATTAGTTGAAAGAGATTCAGTTATCCATCTATAGTGAGTCGTATTAAATT
P57	GGTTAAGTTGGGTATACTTAATGCATAGACTTTCATCCTTAGCACTTCTTCGCGTATTTTAACCTATAGTGAGTCGTATTAAAT	P135	ATTGTTTCGTGATAATGTCTCTTATTAGTTGAAAGAGATTCAATTATCCATCTATAGTGAGTCGTATTAAATT
P58	GGTTAAGTTGGGTATACTTAATGGGCATAGACTTTCATCCTTAGCCCACTTCTTCGCGTATTTTAACCTATAGTGAGTCGTATTAAATT	P162	GATAATGTCGCATAGACTTTCATCCTTAGCGATTCAGTTATCTATAGTGAGTCGTATTAAATT

	Fluorescence labeled DNA oligonucleotides for TIRFM		
P39	*CY3.5*-TTTTGGTTTTTCCTTTGCCGCC	P40	*ATTO488*-GTAGTTTATTTAGCACAGACAAGG

Table 2.4: Oligonucleotides. *Italic* labels indicate fluorophores attached to the respective primers. [#] Source: Dr. Roland Heym

2.5 RNA sequences

RNA (bases relative to start codon AUG)	Short name (used in text)	Sequence (5'→ 3')	Produced by
E3 (1774-1821)		ggGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACGAAACAAU	Dr. Roland Heym
E3 (1771-1814)	**Δ1**	gggAUGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACG	Dr. Roland Heym
E3 (1771-1814)	**Δ2**	gggAUGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACG	Dr. Roland Heym
E3 (1771-1811)	**Δ3**	gggAUGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUC	Dr. Roland Heym
E3 (1771-1821)Δ(1786-1802)GAAA	**E3 (38 nt-loop)**	gAUGGAUAACUGAAUCgaaaGACAUUAUCACGAAACAAU	Dr. Roland Heym
ASH1 E3 (1771-1821) U1780A	**M1**	gAUGGAUAAC**A**GAAUCUCUUUCAACUAAUAAGAGACAUUAUCACGAAACAAU	*In vitro* transcription. Template P45/P130
ASH1 E3 (1771-1821) U1780C	**M2**	gAUGGAUAAC**C**GAAUCUCUUUCAACUAAUAAGAGACAUUAUCACGAAACAAU	*In vitro* transcription. Template P45/P131
E3 (1771-1821) U1780A, G1781C, C1805G, A1806U	**M3**	gggAUGGAUAAC**AC**AAUCUCUUUCAACUAAUAAGAGA**GU**UUAUCACGAAACAAU	Dr. Roland Heym
E3 (1771-1821) C1779G	**M4**	gggAUGGAUAA**G**UGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACGAAACAAU	Dr. Roland Heym
ASH1 E3 (1771-1821) C1779U	**M5**	gAUGGAUAA**U**UGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACGAAACAAU	*In vitro* transcription. Template P45/P135
ASH1 E3 (1771-1821) C1813G	**M6**	gAUGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCA**G**GAAACAAU	*In vitro* transcription. Template P45/P133
ASH1 E3 (1771-1821) C1813U	**M7**	gAUGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCA**U**GAAACAAU	*In vitro* transcription: Template P45/P134
ASH1 E2B (1283-1303)	**E2B (21 nt)**	CCACACCGACGAAAAGUGGCA	Chemical synthesis (Dharmacon)
ASH1 E3 (1774-1811)Δ(1786-1802)GAAA	**E3 (25 nt-loop)**	GAUAACUGAAUCgaaaGACAUUAUC	Chemical synthesis (Dharmacon)
ASH1 E3 (1774-1814)Δ(1786-1802)GAAA	**E3 (28 nt-loop)**	GAUAACUGAAUCgaaaGACAUUAUCACG	Chemical synthesis, in vitro transcription. Template P45/P96
ASH1 E3 (1774-1815)Δ(1786-1802)GAAA	**E3-A (29 nt)**	GAUAACUGAAUCgaaaGACAUUAUCACGA	Chemical synthesis (IBA)
ASH1 E3 (1774-1814)UΔ(1786-1802)_GAAA	**E3-B (29 nt)**	GAUAACUGAAUCgaaaGACAUUAUCACGu	Chemical synthesis (IBA)
ASH1 E3 (1774-1816)Δ(1786-1802)GAAA	**E3-C (30 nt)**	GAUAACUGAAUCgaaaGACAUUAUCACGAA	Chemical synthesis (IBA)
ASH1 E3 (1774-1817)Δ(1786-1802)GAAA	**E3-D (31 nt)**	GAUAACUGAAUCgaaaGACAUUAUCACGAAA	Chemical synthesis (IBA)
ASH1 E3 (1774-1814)	**E3 (41 nt)**	GAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACG	Chemical synthesis (Dharmacon)
ASH1 E3 (1771-1821)	**E3 (51 nt)**	gAUGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACGAAACAAU	*In vitro* transcription. Template P45/P132
*WSC2*_N (430-465)	**_WSC2_N (36 nt)**	AGUUCAAAAACGUCCACGAAAUUGGACACGAAAACU	Chemical synthesis (Dharmacon)
		RNAs with heterologous donor-acceptor stems	
ASH1 E1 (630-687)Δ(651-665)_17nt receptor/GAAA acceptor stem	**E1_1 (64 nt-TL/TLR)**	gGUUAAAAUACGCGAAGAAGUgcuaaggaugaaagucuaugcAUUAAGUAUACCCAACUUAACc	*In vitro* transcription: Template P45/P57
ASH1 E1 (630-687)Δ(651-665)_21nt receptor/GAAA acceptor stem	**E1_2 (68 nt-TL/TLR)**	gGUUAAAAUACGCGAAGAAGUgggcuaaggaugaaagucuaugcccAUUAAGUAUACCCAACUUAACc	*In vitro* transcription: Template P45/P58
ASH1 E2A (1118-1178)Δ(1139-1155)_17nt receptor/GAAA acceptor stem	**E2A (65 nt-TL/TLR)**	GAGACAUUCUAUCGAACAAUUgcuaaggaugaaagucuaugcAAACCAUCUAAGAACAAAAUCUC	*In vitro* transcription: Template P45/P61
ASH1 E2B (1275-1307)_17nt receptor/GAAA acceptor stem	**E2B_1 (54 nt-TL/TLR)**	gcuaaggaugaaagucuaugcAUCUCCCUCCACACCGACGAAAAGUGGCAAGAU	*In vitro* transcription: Template P45/P46
ASH1 E2B (1283-1301)_17nt receptor/GAAA acceptor stem	**E2B_2 (40 nt-TL/TLR)**	gcuaaggaugaaagucuaugcCCACACCGACGAAAAGUGG	*In vitro* transcription: Template P45/P47

RNA (bases relative to start codon AUG)	Short name (used in text)	Sequence (5'→ 3')	Produced by
ASH1 E2B (1275-1307)_ 19nt receptor/GAAA acceptor stem	**E2B_3 (56 nt-TL/TLR)**	gcuaaggcaugaaagugcuaugcAUCUCCCUCCACACCGACGAAAAGUGGCAAGAU	*In vitro* transcription: Template P45/P62
ASH1 E2B (1275-1307)_ 17nt receptor/GAAA acceptor stem	**E2B_4 (56 nt-TL/TLR)**	gcuaaggaugaaagucuaugccAUCUCCCUCCACACCGACGAAAAGUGGCAAGAUg	*In vitro* transcription: Template P45/P63
ASH1 E3 (1774-1811)Δ(1786-1802)_17nt receptor/GAAA acceptor stem	**E3 (42 nt-TL/TLR)**	GAUAACUGAAUCgcuaaggaugaaagucuaugcGACAUUAUC	Chemical synthesis (Dharmacon)
ASH1 E3 (1774-1814)Δ(1786-1802)_17nt receptor/GAAA acceptor stem	**E3 (45 nt-TL/TLR)**	GAUAACUGAAUCgcuaaggaugaaagucuaugcGACAUUAUCACG	Chemical synthesis (Dharmacon)
EAR1 (1572-1621)_17nt receptor/GAAA acceptor stem	**EAR1 (71 nt- TL/TLR)**	gcuaaggaugaaagucuaugcCGAAGAUGAAAAUGAUCACGAGAGAGAUCCCGAACAAUUUUCAGAAUUUG	*In vitro* transcription: Template P45/P60
WSC2 (424-471)_17nt receptor/GAAA acceptor stem	**WSC2_N (69 nt-TL/TLR)**	gcuaaggaugaaagucuaugcAGUGUUAGUUCAAAAACGUCCACGAAAUUGGACACGAAAACUUCAACU	*In vitro* transcription: Template P45/P59
	RNAs for single-particle motility assays in TIRFM		
IST2 (2718-2799) + 22 nt 3' overhang	**IST2**	ggGAUGCUGGAGUGAAGAAUGUCACGAACAAUUCUAAGACAACCGAAUCCUCUUCCUCUUCAUC<u>GGCGGCAAAGGAAAAACCAAAA</u>	*In vitro* transcription. Template produced by PCR with primers P31/P32 on RHP149
ASH1 (1758-1858) + 24-nt 3' overhang	**E3 (77 nt)**	gggAGAGAAUUGAUACAUGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACGAAACAAUUGUACAUUUCUCU<u>CCUUGUCUGUGCUAAAUAAACUAC</u>	*In vitro* transcription. Template produced by PCR with primers P33/P34 on RHP166
2x-ASH1 (1758-1834)-tRNA + 24-nt 3' overhang	**2x E3 (77 nt)-tRNA**	gGGCUACGUAGCUCAGUUGGUUAGAGCAGCGGCCGCCGUA*acauaauu*AGAGAAUUGAUACAUGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACGAAACAAUUGUACAUUUCUCU*aacauau*AGAGAAUUGAUACAUGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACGAAACAAUUGUACAUUUCUCUAUAGAAGGUGGCCGCGGGUCACAGGUUCGAAUCCCGUCGUAGCCACCA<u>CCUUGUCUGUGCUAAAUAAACUAC</u>	*In vitro* transcription. Template produced by PCR with primers P37/P38 on RHP145
ASH1 3'UTR	**ASH1-3'**	GGGCGAAUUGGGUACCGGGCCCCCCCUCGAG<u>CGAACGCAACUGCGAAAUUGAAGG</u>GUACCGUUGCUUAUUUUGUAAUUACAUAACUGAGACAGUAGAGAAUUGAUACAUGGAUAACUGAAUCUCUUUCAACUAAUAAGAGACAUUAUCACGAAACAAUUGUACAUUUCUCUCCUUGUCUGUGCUAAAUAAACUACAAAUAAAAAUUCAUUUUCUUU<u>CAUUUUCAUUCGACACGCCGCGG</u>CAUCUUCGUUUUCUUCACCGAUUAAUUUUAAUCUG<u>CGCCGUGCAGAUAUUUUCGCCC</u>AUUUCGAAGUGAAAUCAUUCUAAUAACUUCUCCUGCGCAGGAUC	*In vitro* transcription. Template produced by PCR with primers P35/P36 on RHP166

Table 2.5: RNA sequences. Small letters indicate heterologous nucleotides, not naturally belonging to the RNA. Bold letters mark introduced base mutations compared to wild type *ASH1*. Underlined letters mark hybridization site for fluorescently labeled DNA-oligonucleotide binding. Italic sequences mark three-way junction sequence from the Twort intron [175].

2.6   Media, plates and antibiotics

Name	Composition	Cell type
LB$_0$ (Luria Bertani):	10 g/l BD Bacto^{TM} tryptone, 5 g/l BD Bacto^{TM} yeast extract, 5 g/l NaCl, pH 7.0 with NaOH, add 15 g/l agar for plates	*E.coli*
HyClone$^®$ SFX-Insect MP^{TM} Cell Culture Medium (Thermo Scientific)	42.32 g/l HyClone$^®$ SFX-Insect MP^{TM}, 0.35 g/l Na bicarbonate, 1.5 ml/l glycerol, pH 6.1-6.4 with NaOH	High five insect cells
SelenoMet^{TM} Medium Base (Molecular Dimensions)	21.6 g in 1 l ddH$_2$O	*E. coli* B834 (DE3)
SF-900^{TM} III SFM (Thermo Scientific)	Serum-free, protein-free, animal origin-free insect cell culture medium	*SF*21 insect cells
SOC	20 g/l BD Bacto^{TM} tryptone, 5 g/l BD Bacto^{TM} yeast extract, 20 mM glucose, 20 mM MgSO$_4$, 10 mM NaCl, 10 mM MgCl$_2$, 2.5 mM KCl, pH 7.0 with NaOH	*E.coli* DH10Bac
SOB	20 g/l BD Bacto^{TM} tryptone, 5 g/l BD Bacto^{TM} yeast extract, 10 mM NaCl, 2.5 mM KCl, pH 7.0 with NaOH	*E.coli*
ZYM 5052 auto-induction media	10 g/l BD Bacto^{TM} tryptone, 5 g/l BD Bacto^{TM} yeast extract	*E.coli* auto-induction

Table 2.6: Media

Name	Stock concentration	Final concentration	Application
Ampicillin	100 mg/ml	100 µg/ml	*E. coli* selection
Carbenicillin	100 mg/ml	100 µg/ml	Selection plates for *E. coli* containing pOPIN vectors
Chloramphenicol	34 mg/ml in absolute ethanol	34 µg/ml	*E. coli* selection
Gentamicin	7 mg/ml	7 µg/ml	*E. coli* DH10Bac selection
Gentamicin	50 mg/ml	10 µg/ml	*Sf*21, High five cells
Glucose	20 % (v/v)	1 % (v/v)	Selection plates for *E. coli* containing pOPIN vectors
IPTG	1 M	40 µg/ml	*E. coli* DH10Bac selection plates
Kanamycin	50 mg/ml	50 µg/ml	*E. coli* (DH10Bac) selection
Tetracyclin	10 mg/ml	10 µg/ml	*E. coli* DH10Bac selection
X-gal	100 mg/ml in DMSO	100 µg/ml	*E. coli* DH10Bac selection plates

Table 2.7: Antibiotics and agar plate supplements

2.7 General buffers and stock solutions

Name	Composition	Application
Buffers for molecular biology		
Denaturing RNA loading dye	2 mM Tris/Cl pH 7.5, 20 mM EDTA, 8 M Urea, 0.025 % (w/v) bromphenol blue, 0.025 % (w/v) xylene cyanol	Denaturing RNA PAGE
DNA-loading dye (6 x)	30 % (w/v) glycerol, 0.25 % (w/v) bromphenol blue	DNA agarose gel electrophoresis
Formamide RNA-loading dye (2 x)	95 % (v/v) formamide, 5 mM EDTA, 0.025 % (w/v) bromphenol blue, 0.025 % (w/v) xylene cyanol, 0.025 % (w/v) SDS	Denaturing RNA PAGE
Native RNA-loading dye (6 x)	60 mM Tris/Cl pH 7.5, 60 % (v/v) glycerol	Native RNA PAGE
SDS-loading dye (4 x)	110 mM Tris/Cl pH 6.8, 40 % (v/v) glycerine, 40 mM DTT, 4 % (w/v) SDS, 0.25 % (w/v) bromphenol blue	SDS-PAGE
Separating gel buffer (4 x)	3 M Tris/Cl pH 8.5, 0.4 % (w/v) SDS	SDS-PAGE

Name	Composition	Application
Stacking gel buffer (4 x)	0.5 M Tris/Cl pH 6.8, 0.4 % (w/v) SDS	SDS-PAGE
TAE (50 x)	2 M Tris base, 50 mM EDTA pH 8.3	DNA agarose gel electrophoresis
TBE (10 x)	765 mM Tris base, 890 mM boric acid, 32 mM EDTA	RNA agarose gel electrophoresis/ PAGE
Protein chemical buffers		
CC-buffer	10 mM Hepes/KOH pH 6.7, 250 mM KCl, 55 mM $MnCl_2$, 15 mM $CaCl_2$,	Production of chemically competent bacteria
EMSA buffer	20 mM Hepes/NaOH pH 7.8, 200 mM NaCl, 4 % (v/v) glycerol, 2 mM $MgCl_2$, (2 mM DTT)*	EMSA
Fixing solution	30 % (v/v) acetic acid, 10 % (v/v) methanol	EMSA
TGS (10 x)	0.5 M Tris, 1.9 M glycine, 1 % (w/v) SDS	SDS-PAGE running buffer
Supplements		
5052 sugar mix (50 x)	25 % (v/v) glycerol, 2.5 % (w/v) glucose, 10 % (w/v) lactose	Auto-induction media
M salt mix (25 x)	1.25 M Na_2HPO_4, 1.25 M KH_2PO_4, 2.5 M NH_4Cl, 0.25 M Na_2SO_4	Auto-induction media
SelenoMet^TM Nutrient Mix (Molecular Dimensions)	5.1 g in 50 ml ddH_2O	SelenoMet^TM Medium Base
Trace metal mix (1000 x)	40 mM HCl, 50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$, 10 mM $ZnCl_2$, 2 mM $CoCl_2$, 2 mM $CuCl_2$, 2 mM Na_2MoO_4, 2 mM H_3BO_3	Auto-induction media

Table 2.8: General buffers and stock solution. *In absence of cysteins, DTT was not added.

2.8 Molecular Biology

For all molecular biology techniques and cultivation of organisms, respectively, sterile instruments and solutions were used.

2.8.1 Preparation of chemically competent *E.coli*

Chemically competent *E.coli* were prepared based on [176]. In brief, 200 ml SOB media (Table 2.6) containing the appropriate antibiotic(s) (Table 2.7) for the resistance genes of the respective strain (Table 2.1) were inoculated with 10 ml bacterial overnight pre-culture and incubated at 30 °C, 160 rpm until an $OD_{600\,nm}$ of 0.4-0.6 was reached. Bacterial growth and cell density were monitored

with a UV-1600 PC UV/Vis spectrophotometer (VWR). After cold shock on ice for at least 10 min, the cell suspension was centrifuged for 10 min, 5,900 x g, 4 °C. Harvested cells were washed with 40 ml CC-buffer (Table 2.8) and the cell pellet gently dissolved in 7.44 ml CC-buffer supplemented with 0.56 ml DMSO. After chilling on ice for 10 min, 100 µl aliquots were flash frozen in liquid nitrogen and stored at -80 °C.

2.8.2 Transformation of competent *E. coli* cells

For transformation of chemically competent bacteria 100-500 ng of plasmid DNA was added to 100 µl *E. coli* cells and incubated on ice for 5 min. Cells were heat shocked for 45 s at 42 °C and placed back on ice for 2 min. 800 µl LB_0, or SOC medium in case of *E. coli* DH10Bac, were added and the cells incubated at 37 °C for 45 min or 4 h, respectively, to let the resistance genes be transcribed and the antibiotics-converting enzymes be synthesized. Cells were pelleted at 900 x g for 2 min, plated onto selection media plates and incubated o.n. at 37 °C. DH10Bac cells grew for 48 h at 37 °C until blue and white colonies appeared. To avoid false-positive clones, double consecutive selection of one white colony was performed.

2.8.3 Amplification and storage of plasmid DNA

For amplification of plasmid DNA, single colonies were transferred into 5 ml of LB containing the corresponding antibiotics and grown in a shaker at 150 rpm, 37 °C, o.n.. Cells were harvested and DNA extracted with the NucleoSpin® Plasmid kit (Machery-Nagel) according to manufacturers' guidelines. Plasmids were stored at -20 °C.

In case of bacmid DNA, NucleoSpin® Plasmid kit ingredients (Machery-Nagel) were used for cell lysis. To protect the integrity of these large plasmids, DNA was precipitated by the addition of 0.8 ml isopropanol. After chilling on ice for 10 min, DNA was pelleted by centrifugation for 30 min at 16.100 x g and washed with 500 µl of 70 % (v/v) ethanol. The air-dried DNA pellet was gently dissolved in sterile kit elution buffer and stored at 4 °C.

2.8.4 Molecular cloning

Standard molecular cloning methods were performed as described in reference [177]. Polymerase chain reaction (PCR) was used to amplify specific DNA sequences of template plasmids. Salt-free DNA oligonucleotides serving as primers were ordered at Eurofins MWG Biotech and are listed in Table 2.4. PCR products were analyzed by agarose gel electrophoresis as described in section 2.8.5 and purified using the Nucleospin® Gel and PCR Clean-up kit (Macherey-Nagel). Double-stranded plasmid and insert DNA were digested with adequate restriction enzymes (New England BioLabs). Cleaved vector DNA was separated by means of an agarose gel and both, vector and insert purified with the Nucleospin® Gel and PCR Clean-up kit (Macherey-Nagel). Linearized vector backbone was 5' dephosphorylated with FastAP^{TM} Thermosensitive Alkaline Phosphatase (Thermo Fisher) and ligated with PCR insert using T4 DNA Ligase (Thermo Fisher). Enzymes and corresponding buffers were applied according to manufacturer's instructions. 10 µl of ligated sample was transformed in chemically competent *E.coli* XL1 Blue cells as described in section 2.8.2.

Point mutations and amino-acid deletions were either introduced by using the QuickChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies) or by three-point PCR using the overlap-extension technique [178].

Plasmids based on pOPIN vectors were produced using the In-Fusion® HD Cloning Kit (Clontech Laboratories) according to manufacturer's guidelines. Correctness of cloned plasmids was verified by DNA sequencing at Eurofins MWG Biotech. Applied primers are listed in Table 2.4, cloning strategies and generated plasmids are summarized in Table 2.3.

2.8.5 Agarose gel electrophoresis

Depending on the size of the DNA, separation was carried out by using 0.5-1.5 % (w/v) agarose (Biozym) supplemented with GelRed^{TM} (Biotium) 1:10,000 in 1 x TAE buffer (Table 2.8). Samples were supplemented with 1 x DNA-loading dye (Table 2.8) and electrophoresis set up with a constant voltage of 120 V using 1 x TAE buffer. Fluorescence was visualized by UV-light using the Fusion SL imaging system (Vilber Lourmat) device. As marker for DNA sizing 10 µL of either GeneRuler^{TM} 1 kb or 100 bp DNA Ladder (Thermo Scientific) were used.

In case of RNA 1 x TBE buffer (Table 2.8) was used for gel preparation and as running buffer. Unless stated otherwise, native RNA samples were unfolded for 10 min, 70 °C, refolded via snap cooling on

ice and supplemented with native RNA-loading dye (Table 2.8). To assess RNA integrity, samples were denatured with 1 x formamide RNA-loading dye for 2 min, 90 °C, and both, native and denatured samples separated at 80 V and 120 V const., respectively.

2.8.6   Handling of RNA

For work with RNA in general, RNase-free materials, reagents and RNase-free buffers containing diethylpyrocarbonate (DEPC)-treated H_2O, were used. RNase-free water was prepared by dissolving 0.05 % (v/v) DEPC (Roth) via stirring in ddH_2O and incubating it o.n. at 37 °C. DEPC was depleted by autoclaving. Surface decontamination was achieved by using RNase AWAY® (Fisher Scientific).

RNAs (sequences see Table 2.5) were either produced by *in vitro* transcription (see section 2.10.4 and 2.10.5) or purchased chemically synthesized and polyacrylamide gel electrophoresis (PAGE)-purified at IBA Lifesciences or Dharmacon™. 2'-ACE protected RNA from Dharmacon™ was deprotected by dissolving the RNA in 400 µl 2'-deprotection buffer (100 mM acetic acid pH 3.8 with TEMED) and incubation for 30 min, 60 °C. RNA was lyophilized and stored at -20 °C.

2.9   Protein expression and purification

2.9.1   Protein expression in bacteria

A preparatory culture of 50 ml was inoculated with several colonies from freshly transformed *E. coli* cells (see section 2.8.2) and incubated at 37 °C, 150 rpm until the suspension was dense. Large-scale cultivation comprising 3 l LB medium containing respective antibiotics was set up. In case of pETM or pOPIN-M vectors 0.2 % (w/v) glucose was added. After inoculation and growth at 37 °C to an OD_{600} of 0.3, cells were cooled to 18 °C within 1-1.5 h and when reaching an OD_{600} of 0.6-0.8, 0.25 mM IPTG was added. Target genes were expressed at 18 °C, 150 rpm, for 17-20 h.

For the expression of T7 RNA polymerase present in a pQE vector in *E.coli* M15 DZ291 (gift from Dr. Arie Geerlof, Protein Expression and Purification Facility, Helmholtz-Zentrum München), the dense 50 ml preculture was spun down 900 x g, 5 min, RT and resuspended in 1.5 l fresh LB medium containing the respective antibiotics. IPTG induction and target gene expression was performed as described above.

For the production of L-selenomethionine (SeMet)-containing proteins the methionine auxotrophic *E. coli* B834 (DE3) strain was used. SelenoMet™ Medium (Molecular Dimensions) was prepared according to manufacturer's instructions and supplemented with 100 mg/l SeMet. Three 100 ml pre-cultures were grown o.n. at 37 °C, 150 rpm and transferred to 3 l pre-warmed SelenoMet™ Medium. After 12 h growth at 37 °C, 150 rpm, bacteria had an OD_{600} of ~ 0.6.

Protein production by auto-induction was performed as essentially described in [179]. 100 ml LB medium from an o.n. preculture was added to 3 l pre-warmed ZYM 5052 auto-induction media (Table 2.6), supplemented with 2 mM $MgSO_4$, 2 x trace metal mix, 1 x 5052 sugar mix, 1 x M salt mix (Table 2.8) and appropriate antibiotics (Table 2.7). Bacteria were grown for 3 h at 37 °C, 150 rpm. Expression of target genes was automatically induced and proteins were expressed at 18 °C, 150 rpm, for 24-30 h.

Cells were harvested by centrifugation, 10 min, 2,449 x g, 4 ° C, resuspended in 25 ml appropriate lysis buffer including one cOmplete™, EDTA-free Protease Inhibitor Cocktail tablet (Roche Diagnostics) and either directly used for protein purification or flash frozen in liquid nitrogen and stored at -80 °C.

2.9.2 Production of baculovirus and protein expression in insect cells

For the expression of recombinant proteins in insect cells the Bac-to-Bac® Baculovirus Expression System (Invitrogen) was essentially applied as described in the manufacturer's manual. For each transfection reaction 2 ml *Sf*21 cells with a density of 0.4 x 10^6 cells/ml were seeded in a 6-well plate and allowed to adhere for 30-45 min at 27.5 °C. The components of the transfection mix were 1.5 µg DNA and an adjusted amount of transfection reagent depending on the plasmid. In case of pFBDual plasmids 3 µl FuGENE® HD Transfection reagent (Promega) were used. For pOPIN vectors 6 µl Cellfectin® Reagent (Invitrogen) and 0.5 µg Bac10:KO_{1629} [170] (linearized with Bsu36I) were added. The ingredients were mixed in 200 µl SF-900™ III SFM medium (Thermo Scientific) and incubated for 30 min at room temperature. Per transfection reaction, 200 µl transfection mix was added and the cells incubated for 4 d at 27.5 °C. The supernatant of two wells containing the baculovirus was added to 10 ml *Sf*21 cells (1.4 x 10^6 cells/ml) and incubated for 3-4 d at 27.5 °C, 85 rpm until signs of infection were visible. The P1 viral stock was separated from cells and debris by centrifugation for 10 min, 2,000 x g and filtered using a 0.22 µm filter (Merck Millipore). To obtain highest viral titers,

the virus was further amplified in 250 ml Sf21 cell culture (0.4 x 10^6 cells/ml) inoculated with 1.5 ml P1 virus. After incubation for 3-4 d at 27.5 °C, 85 rpm, the resulting P2 viral stock was harvested as described for P1 and stored at 4 °C.

The production of recombinant protein was started by infecting 700 ml High Five cells (1 x 10^6 cells/ml) with 10 ml P2 virus stock and the proteins expressed for 60-70 h at 27.5 °C, 85 rpm. Insect cells were harvested and stored as described for bacterial cells (see section 2.9.1), only differing in centrifugation speed (2,000 x g).

2.9.3 Cell lysis

Unless stated otherwise, harvested bacterial cells (see 2.9.1) were lysed at 4 °C by sonication with a Branson sonifier 250 (Emerson) 3 x 4 min, amplitude 30 %, output 6. Insect cells (see 2.9.2) were lysed by sonication 3 x 2 min, amplitude 40 %, output 5. Soluble protein was separated from insoluble fragments by centrifugation for ≥ 30 min, at 39,191 x g, 4 °C. High Five lysate was centrifuged twice and the supernatant filtered with a 2.7 µm filter (Whatman).

2.9.4 Purification of T7 RNA polymerase

The purification of T7 RNA polymerase was performed in accordance with a purification protocol from Dr. Arie Geerlof (Protein Expression and Purification Facility, Helmholtz-Zentrum München). Harvested cells (see 2.9.1) were lysed by sonication for 2 x 3 min, amplitude 30 %, output 6, in 40 ml lysis buffer containing 50 mM Tris/Cl pH 8.0, 300 mM NaCl, 20 mM imidazole, 10 mM MgSO$_4$, 10 µg/ml DNase I, 1 mM AEBSF/Cl, 1 mM β-mercaptoethanol and 1 mg/ml lysozyme. Cleared lysate was filtered with a 0.22 µm filter (Merck Millipore) and loaded on a 5 ml HisTrap HP column (GE Healthcare), pre-equilibrated with Ni-binding buffer (50 mM Tris/Cl pH 8.0, 300 mM NaCl, 20 mM imidazole, 1 mM β-mercaptoethanol). The column was extensively washed with Ni buffer containing 50 mM imidazole and the protein eluted with Ni-elution buffer (Ni-binding buffer with 300 mM imidazole). Protein containing fractions were dialyzed o.n. against HyAP-binding buffer (20 mM K$_2$HPO$_4$/KH$_2$PO$_4$ pH 7.5, 50 mM KCl, 0.05 mM EDTA, 3 mM β-mercaptoethanol) and applied to a pre-equilibrated 15 ml BioGel® HT column (BioRad). After extensive washing with HyAP-binding buffer, bound protein was eluted with HyAP-elution buffer (HyAP-binding buffer supplemented with 568 mM (NH$_4$)$_2$SO$_4$) and concentrated via dialysis o.n. in storage buffer (20 mM K$_2$HPO$_4$/KH$_2$PO$_4$ pH 7.5,

100 mM NaCl, 1 mM EDTA, 1 mM DTT, 50 % (v/v) glycerol). 500 µl aliquots à 10 mg/ml were flash-frozen in liquid nitrogen and stored at -80 °C.

2.9.5 Purification of HRV-3C protease

The purification of His$_6$-GST-HRV-3C protease was performed as essentially described in a single-step purification protocol from Dr. Arie Geerlof (Protein Expression and Purification Facility, Helmholtz-Zentrum München). Harvested *E. coli* BL21 Gold (DE3) pRARE cells containing the pGEX/3C plasmid from 1 l auto-induction media (see section 2.9.1) were lysed by sonication for 2 x 3 min, amplitude 30 %, output 6, in 40 ml lysis buffer containing 50 mM Tris/Cl pH 8.0, 300 mM NaCl, 10 % (v/v) glycerol, 0.2 % (v/v) NP-40, 0.02 % (v/v) 1-thioglycerol, 1 mg/ml lysozyme and 2 µg/ml DNase I. The cleared supernatant was filtered with a 0.22 µm filter (Merck Millipore) and loaded on a 5 ml GSTrap™ FF column (GE Healthcare), pre-equilibrated with wash buffer (50 mM Tris/Cl pH 8.0, 300 mM NaCl, 10 mM reduced gluthathione, 10 % (v/v) glycerol, 0.01 % (v/v) 1-thioglycerol). After extensively washing the column for the removal of all unspecifically bound protein, 3C-protease was eluted with elution buffer (wash buffer supplemented with 10 mM reduced gluthathione) and protein containing fractions dialyzed in storage buffer (50 mM Tris/Cl pH 8.0, 150 mM NaCl, 10 mM EDTA, 20 % (v/v) glycerol, 0.01 % (v/v) 1-thioglycerol), o.n., 4 °C. 50 µl aliquots à 2 mg/ml were flash-frozen and stored at -80 °C.

2.9.6 Purification of She2p and its mutants

SHE2 wild type and its mutants (*SHE2(6-240 C-S), SHE2(6-243 C-S), SHE2(6-246), SHE2(6-246 C-S), SHE2_N36S_ R63K,* and *SHE2_E172A_F176A),* present in pGEX-6P-1 were transformed in *E. coli* BL21 (DE3) Star cells. Protein expression was achieved by auto-induction (see section 2.9.1). Purification was done as essentially described in references [139, 150]. After cell lysis and centrifugation (see capter 2.9.3), soluble protein present in GST-A1 buffer (20 mM Hepes/NaOH pH 7.5, 500 mM NaCl, 1 mM EDTA, 2 mM DTT and 1 mM PMSF) was loaded on a GSTrap FF (5 ml) column (GE Healthcare), pre-equilibrated with GST-A1 buffer. Bound protein was extensively washed with each ≥ 10 CV GST-A1 buffer, GST-A2 (GST-A1 including 1 M NaCl) and Q-A buffer (10 mM Hepes/NaOH pH 7.9, 150 mM NaCl, 1 mM EDTA, 2 mM DTT). In case of GST-tag removal o.n., 4 °C, 50 µg HRV-3C-protease in Q-A buffer was applied to the column. His$_6$-GFP-tagged variants of She2p wild type, She2p (N36S, R63K)

and She2p Δhelix E were initially captured via HisTrap FF 5 ml column (GE Healthcare) in Ni-A1 (20 mM Hepes /NaOH pH 7.5, 500 mM NaCl, 20 mM imidazole) buffer and washed, each ≥ 10 CV with Ni-A1 buffer, Ni-A2 buffer (Ni-A1 including 1 M NaCl) and 20 % Ni-B (Ni-A1 including 100 mM imidazole). Bound proteins were eluted by gradient elution over 20 CV 10 % - 100 % Ni-B and dialysed in Q-A buffer, o.n., 4 °C. The cleaved or His$_6$-GFP-tagged protein was loaded onto a HiTrap Q FF (5 ml) column (GE Healthcare). After gradient elution over 20 CV 0 % - 100 % Q-B buffer (Q-A including 1 M NaCl) purest fractions were concentrated in an Amicon® Ultra Centrifucal Filter, MWCO 30,000 (Merck Millipore) 10 min, 4409 x g, 4 °C, spun down 10 min, 16,100 x g, 4 °C and loaded on a Superdex 200 16/60 GL column (GE Healthcare) in the final SEC buffer (20 mM Hepes/NaOH pH 7.8, 200 mM NaCl and 2 mM DTT in case of cysteins being present in the protein).

2.9.7   Purification of She3p

Full-length She3p-His$_6$ was coexpressed with She2p wild type from baculovirus-infected insect cells (see section 2.9.2) and purified by multistep procedures as essentially described in [139]. After cell disruption (see section 2.9.3), the supernatant in Ni-A1 buffer (20 mM Hepes/NaOH, pH 7.8, 500 mM NaCl, and 20 mM imidazole, 1 mM PMSF) was loaded on a 5 ml HisTrap FF column (GE Healthcare). The bound protein was first washed with Ni-A2 buffer (20 mM Hepes/NaOH, pH 7.8, 1 M NaCl, and 40 mM imidazole) and second with Hep-A buffer (20 mM Hepes/NaOH, pH 7.8, and 200 mM NaCl). Elution was achieved by supplementing buffer Hep-A with 300 mM imidazole. Protein passed a 5 ml HiTrap Q FF column (GE Healthcare) before it was bound on a 5 ml Heparin HP column (HiTrap; GE Healthcare). Here, it was washed with 10 % Hep-B buffer (20 mM Hepes/NaOH, pH 7.8, 1 M NaCl, 2 mM EDTA, and 2 mM DTT) and eluted with Hep-B buffer. Finally, She3p was loaded on a Superose 6 10/300 GL column (GE Healthcare) in a buffer that contained 20 mM Hepes/NaOH, pH 7.8, 500 mM NaCl, and 2 mM DTT.

For motility experiments, the purification of full-length She3p was extended with an additional competition step to completely remove remnants of wild-type She2p. Hence, She3p bound to the 5 ml HisTrap FF column was incubated with 30 μM She2p (N36S, R63K) for 15 min at 4 °C during the purification and replacing She2p being capable of binding to RNA.

2.9.8   Purification of She3p C-terminal fragments

Unless stated otherwise She3p C-terminal fragments were expressed by IPTG induction (see section 2.9.1) in *E. coli* BL21(DE3)Star and the cells lysed as described in section 2.9.3.

GST-tagged She3p fragments (RHP85, RHP94, FE#4, FE#9, #26, #28, see Table 2.3) were purified essentially as described for She2p (see section 2.9.6). Due to low yields, She3p was directly applied to gel filtration after HRV-3C-protease cleavage using a Superdex 75 10/300 GL column (GE Healthcare) in SEC buffer (20 mM Hepes/NaOH pH 7.8, 200 mM NaCl). Concentrations were estimated by comparing band intensities on SDS-gels with known concentrations of She2p (6-240, C-S) and the protein was stored at -20 °C.

MBP-tagged She3p fragments (RHP143, FE#6, FE#16-#18, see Table 2.3) present in MBP-A buffer (20 mM Hepes/NaOH pH 7.5, 200 mM NaCl, 5 mM EDTA) were applied to a self-packed 30 ml amylose column (BioRad) and washed with $\geq$ 5 CV MBP-A and $\geq$ 25 CV MBP-B buffer (20 mM Hepes/NaOH pH 7.8, 200 mM NaCl) before they were eluted with MBP-C buffer (20 mM Hepes/NaOH pH 7.8, 500 mM NaCl, 10 mM maltose). The tag was cleaved from the protein with 50 µg HRV-3C-protease in an appropriate dialysis membrane (Roth) in 2 l of SP-A buffer containing 20 mM Hepes/NaOH pH 7.8, 150 mM NaCl, or in case of C-terminally tagged She3p-His$_6$ fragments in Ni-A buffer (20 mM Hepes/NaOH pH 7.8, 500 mM NaCl, 20 mM imidazole) o.n., 4 °C. She3p was either captured via a 5 ml HiTrap SP FF column (GE Healthcare), washed and gradient-eluted (0 % - 100 % SP-B (SP-A including 1 M NaCl), or applied to HisTrap FF 5 ml column (GE Healthcare) and eluted with Ni-A buffer containing 500 mM imidazol, respectively. Finally the protein was purified with a Superdex 75 10/300 GL column (GE Healthcare) in the final SEC buffer (20 mM Hepes/NaOH pH 7.8, 200 mM NaCl). Concentrations were estimated by comparing band intensities on SDS-gels with known concentrations of She2p (6-240, C-S) and the protein was stored at -20 ° C.

Cells containing His$_6$-SUMO-She3p constructs (FE#89, #94, #98 and #99, #108-#110, #113 and #116, see Table 2.3) were lysed as described above. Proteins in Ni-A1 buffer were loaded on a pre-equilibrated HisTrap FF 5 ml column (GE Healthcare) and washed $\geq$ 5 CV Ni-A1 buffer, $\geq$ 10 CV Ni-A2 buffer (Ni-A1 including 1 M NaCl) and $\geq$ 5 CV 10 % Ni-B (Ni-A1 including 300 mM imidazole). Bound proteins were eluted by gradient elution over 20 CV 10 % - 100 % Ni-B. In case of His$_6$-SUMO-She3p(331-425)His$_6$ and His$_6$-SUMO-She3p(331-405), which were expressed in *E. coli* BL21 Gold (DE3) pRARE, tag-cleavage with 50 µg HRV-3C-protease took place o.n., at 4 °C in Hep-A buffer (10 mM Hepes/NaOH pH 7.9, 150 mM NaCl, 2 mM DTT) and SEC buffer, respectively. She3p(331-425)His$_6$ was

separated from its tag via an additional HiTrap™ Heparin HP (5 ml) (GE Healthcare) purification step using a gradient elution over 30 CV 0 % - 100 % Hep-B (Hep-A including 300 mM NaCl). To separate She3p(331-405) from its His_6-SUMO tag, a second HisTrap capture step was added, which was a gradient ranging from 0 % SEC buffer - 100 % Ni-B in 30 CV. As described before, purest fractions were concentrated in centrifugal filters with appropriate MWCOs, spun down and finally subjected to size-exclusion chromatography using a Superdex 75 16/60 GL column (GE Healthcare) in SEC buffer without DTT.

2.9.9 Purification of She2p-She3p fusions

pOPIN-J plasmids for She2p-$(GGSGG)_{1/2}$-She3p fusion proteins (see Table 2.3) were transformed in *E.coli* BL21 Gold (DE3) pRARE. Expression of the His_6-GST-tagged protein was done by ITPG induction (see section 2.9.1). Cells were lysed as described for His_6-SUMO-tagged She3p proteins. Purification was carried out as described for She2p, except for GST-A2, Q-A, Q-B and SEC buffer being without DTT. A Superdex 75 16/60 GL column (GE Healthcare) was used for the final purification step.

She2p(6-246, C-S)-GGSGG-She3p(331-343) was also purified from *E.coli* BL21 Gold (DE3) pRARE cells containing the target gene in a pOPIN-M vector (FE#44, see Table 2.3). Cells were lysed by sonication 4 x 4 min, amplitude 40 %, Output 6, spun down and the supernatant sterile filtered. His_6-MBP-tagged fusion protein, present in MBP-A buffer (20 mM Hepes/NaOH pH 7.5, 500 mM NaCl, 2 mM EDTA, 2 mM DTT) was applied to a self-packed 30 ml amylose column and washed with each ≥ 10 CV MBP-A and MBP-B buffer (20 mM Hepes/NaOH pH 7.5, 200 mM NaCl, 2 mM DTT) before it was eluted with MBP-C buffer (20 mM Hepes/NaOH pH 7.9, 500 mM NaCl, 10 mM maltose). The tag was cleaved from the protein with 100 μg 3C protease in an appropriate dialysis membrane (Roth) in 2 l of dialysis buffer containing 20 mM Hepes/NaOH pH 7.9, 500 mM NaCl and 0.1 % (v/v) triton x, o.n., 4 °C. After cleavage the protein was dialysed against Ni-A buffer (10 mM Hepes/NaOH pH 7.9, 150 mM NaCl) and loaded on a HisTrap FF 5 ml column (GE Healthcare). The fusion protein was separated from its tag via gradient elution over 30 CV 0 % - 50 % Ni-B (20 mM Hepes/NaOH pH 7.9, 200 mM NaCl, 500 mM imidazole. To avoid precipitation, purest fractions were slowly concentrated in an Amicon® Ultra Centrifugal Filter, MWCO 30,000 (Merck Millipore) 10 min, 787 x g, 4 °C and subjected to a Superdex 200 16/60 GL column (GE Healthcare) in the final SEC buffer (20 mM Hepes/NaOH pH 7.8, 200 mM NaCl).

2.9.10 Purification of FLAG-Myo4p full-length

FLAG-tagged, full-length yeast Myo4p was coexpressed with She3p, She4p, Cmd1p, and Mlc1p using the transposition system (Bac-to-Bac®; Invitrogen). Purification of Myo4p was performed based on a previously published protocol [164]. Cells were pelleted and resuspended in wash buffer (10 mM imidazole, pH 7.4, 300 mM NaCl, 5 mM MgCl$_2$, and 1 mM EGTA) supplemented by 7 % (w/v) sucrose and protease inhibitors. After sonication, 7 µg/ml yeast Cmd1p, 7 µg/ml yeast Mlc1p, and 2 mM ATP were added, and the lysate was centrifuged. The supernatant was incubated with anti-FLAG M2 affinity gel resin (Sigma-Aldrich) for 1 h at 4°C, and the resin washed with wash buffer. Bound protein was eluted with 0.15 mg/ml FLAG peptide in wash buffer. The eluate was dialyzed in motility buffer (25 mM imidazole, pH 7.4, 50 mM KCl, 4 mM MgCl$_2$, and 1 mM EGTA) supplemented with 2 mM DTT, followed by dialysis in motility buffer supplemented with 2 mM DTT and 50 % (v/v) glycerol. Concentrations were estimated by comparing band intensities on SDS-gels with known concentrations of BSA and the protein was stored at -20 ° C.

2.9.11 Purification of Myo4-C

The C-terminal part of *Myo4-C* (aa 978-1471), present in pGEX-6P-1 was expressed in *E. coli* BL21 (DE3) Star and purified as essentially described in [155]. Cells were lysed (see section 2.9.3) in lysis buffer (20 mM Tris/Cl pH 7.5, 500 mM NaCl, 1 mM EDTA, 1 mM DTT) and soluble proteins loaded on a GSTrap FF 5 ml column (GE Healthcare). In case the GST-tag was cleaved, 50 µg HRV-3C-protease (see section 2.9.5) in buffer AM4 were incubated over night on the column. Cleaved protein was directly washed on a pre-equilibrated HiTrap™ SP 5 ml column (GE Healthcare) with buffer AM4 (10 mM Tris/Cl pH 8.25, 200 mM NaCl, 1 mM DTT) and eluted with a gradient over 20 CV 0 % - 100 % buffer BM4 (buffer AM4 with 1 M NaCl). After concentrating the protein, insoluble aggregates were separated by centrifugation (10 min, 11,000 x g, 4 °C) and the soluble part purified on a Superdex 200 10/300 GL column (GE Healthcare) in 20 mM Hepes/NaOH pH 7.8, 200 mM NaCl, 2 mM DTT. Pure protein fractions were concentrated in centrifugal filters, flash-frozen and stored at -80 °C.

2.9.12 Purification of Cmd1p and Mlc1p

Yeast calmodulin (Cmd1p) and myosin light chain (Mlc1p) were expressed in *E. coli* BL21(DE3) by IPTG induction (see section 2.9.1) and the cells lysed in 50 mM Tris/Cl, pH 7.5, 2 mM EDTA, and 1 mM

PMSF as described in section 2.9.3. The supernatant was boiled for 5 min, cooled, and clarified by centrifugation (39,000 x g, 20 min). For yeast calmodulin, 5 mM $CaCl_2$ and 1 mM DTT were added, and the protein bound to a 5 ml HiTrap Phenyl FF (high sub) column (GE Healthcare) pre-equilibrated with 50 mM Tris/Cl, pH 7.5, 100 mM NaCl, and 5 mM $CaCl_2$. The column was washed with the same buffer containing 0.1 mM $CaCl_2$ and then with the same low-calcium buffer with 0.5 M NaCl. Bound protein was eluted with 50 mM Tris/Cl, pH 7.5, and 1 mM EGTA and finally loaded onto a Superose 12 10/300 GL column (GE Healthcare) in complex buffer (20 mM Hepes/NaOH, pH 7.8, 100 mM NaCl, 25 mM KCl, 2 mM $MgCl_2$, and 1 mM EGTA). For Mlc1p, the clarified supernatant was precipitated with ammonium sulfate to 80 % saturation after boiling, and the protein was pelleted for 20 min at 31,000 x g. The pellet was dissolved in a buffer containing 20 mM Hepes/NaOH, pH 7.4, 100 mM NaCl, and 1 mM EDTA and, as the final purification step, loaded onto a Superdex 200 10/300 GL column (GE Healthcare) in motility buffer (20 mM imidazole, pH 7.4, 50 mM KCl, 4 mM $MgCl_2$, 1 mM EGTA). Concentrations were estimated by comparing band intensities on SDS-gels with known concentrations of BSA.

2.10 Chemical and biophysical methods

2.10.1 Determination of DNA/RNA concentration

DNA and RNA concentration was determined photometrically with a Nanodrop® ND-1000 UV/VIS Spectrophotometer (Peqlab) by measuring $A_{260\,nm}$. Additionally absorption readings at 280 nm were taken to detect impurities from proteins. Purity was acceptable if $A_{260\,nm}/A_{280\,nm}$ was within the range of 1.8-2.0. Concentrations were calculated using the following relationships: $A_{260\,nm}$ of 1.0 = 50 ng/µl dsDNA or 40 ng/µl RNA.

2.10.2 Determination of protein concentration

Protein concentrations were determined by UV spectroscopy using a Nanodrop® ND-1000 UV/VIS Spectrophotometer (Peqlab). According to the Lambert-Beer law $A = \varepsilon \cdot c \cdot d$ with A absorbance, ε molar extinction coefficient [$M^{-1}cm^{-1}$], c protein concentration [M] and d cell length [cm], the measured absorbance at 280 nm is correlated with the molar concentration. Molar extinction coefficients were calculated with the ProtParam tool of the ExPASy server [180]. Proteins were considered to be nucleic-acid-free if $A_{260\,nm}/A_{280\,nm}$ was between 0.5-0.6.

Mass concentrations of peptides or proteins with low molar extinction coefficients were either determined using Direct Detect® Infrared Spectrometer (Merck Millipore) or estimated by comparing protein band intensities on SDS-PAGE gels with similar sized proteins of known concentrations, respectively. These estimations were done in triplicates.

2.10.3 Polyacrylamide gel electrophoresis

Electrophoretic separation of proteins due to their molecular weights was performed as essentially described in [181]. Depending on the size of the protein, 10-19 % polyacrylamide gels of 1.5 mm thickness were used. Separating gels were covered with 5 % stacking gels. To assure complete protein denaturation, samples were supplemented with an adequate amount of SDS-loading dye (Table 2.8) and incubated 5 min at 95 °C. Proteins were separated at a constant voltage of 220 V using 1 x TGS buffer (Table 2.8) and either visualized by coomassie-staining with PageBlue™ Protein Staining Solution (Thermo Scientific) or by silver-staining using the Pierce™ Silver Stain Kit (Thermo Scientific) according to manufacturer's instructions.

Denaturing TBE-PAGE was used to assess the integrity and purity of RNAs. Unless stated otherwise, 10 % polyacrylamide gels containing 8 M urea and 1 x TBE (Table 2.8) were prepared. 250 ng RNA was supplemented with either denaturing RNA- or formamide-loading dye (Table 2.8) and incubated for 1 min at 98 °C. RNA species were separated for 1.5 h at 110 V const. in 1 x TBE.

Native PAGE served to verify for homogenous RNA folding. Unless stated otherwise, the RNA was denatured 10 min at 70 °C, refolded via snap cooling on ice and supplemented with native RNA-loading dye (Table 2.8). RNAs were separated on 20 % native TBE-PAGE gels as described for denaturing TBE-PAGE gels. RNAs from both, native and denaturing TBE-PAGE were stained by shaking the gels for 2 min in 1 x TBE with GelRed™ (Biotium) 1:10,000. After rinsing the gels with 1 x TBE, fluorescent RNA bands were visualized by UV-light using the Fusion SL imaging system (Vilber Lourmat) device.

2.10.4 Small-scale *in vitro* transcription

All RNAs prepared for single-particle motility assays or used in EMSA experiments, except for the E3-(28 nt-loop) construct (Table 2.5), were produced in small-scale *in vitro* transcription preparations.

DNA templates for RNAs used in motility experiments were generated by PCR (section 2.8.4). To produce the at least partially double stranded template DNA for RNAs used in EMSA experiments, 100 pmol of each HPLC-purified forward and reverse primers (Table 2.4) were mixed in annealing buffer (final concentration of 6 mM Tris/Cl pH_{RT} 8.0, 60 mM NaCl, 0.6 mM EDTA), unfolded for 1 min, 98 °C and annealed by cooling to room temperature. 20 pmol template DNA was *in vitro* transcribed and the RNA purified using the MEGAshortscript™ T7 Transcription Kit (Ambion) according to manufacturer's instructions. In brief, after *in vitro* transcription, the remaining DNA was DNase I digested, RNA phenol-chloroform-isoamylalcohol extracted, ethanol precipitated and dissolved in RNase-free water. Depending on the size of the RNA, the integrity was confirmed using a 10 % denaturing TBE-PAGE gel and homogenous folding was assessed on a 10-20 % native TBE-PAGE gel (see section 2.10.3).

2.10.5 Large-scale *in vitro* transcription and RNA purification

To produce large amounts of RNA for crystallization trials, SAXS or NMR experiments, large-scale *in vitro* transcription was used. For DNA template annealing, 4 µM HPLC-purified forward primer (containing the T7 RNA polymerase promoter region) and 3.3 µM HPLC-purified reverse primer (containing the reverse-complement target DNA sequence and T7 RNA polymerase promoter region, Table 2.4) supplemented with an optimized amount of $MgCl_2$ were incubated for 5 min, 60 °C in a total volume of 1.2 ml and cooled to 20 °C. Each 10 ml *in vitro* transcription reaction contained the DNA template mixture, 4 mM of each NTP, 24 mM $MgCl_2$, 80 mg/ml PEG 8000 and 0.5 mg/ml T7 RNA polymerase (see section 2.9.4) in the presence of TRX buffer (40 mM Tris/Cl pH_{RT} 8.1, 1 mM spermidine, 0.1 ‰ triton x-100, 5 mM DTT) and was incubated for 3 h, 37 °C. Precipitants were removed by spinning the sample 10 min, 16,873 x g. Transcribed RNA was precipitated by adding 0.1 V 3 M NaOAc and 3.5 V EtOH and subsequent chilling at -20 °C for ≥ 30 min.

The RNA was pelleted for 10 min, 16,000 x g, 4 °C, solubilized in denaturing RNA loading dye (Table 2.8) and its species separated on a 8 % denaturing TBE-PAGE gels in an Owl™ sequencing chamber (Thermo Fisher Scientific) in 1 x TBE running buffer for 17 h, const. 300 V. Transcripts were visualized by UV shadowing and the desired band excised from the gel. The RNA is subsequently extracted by electroelution using the Whatman® Elutrap electroelution system (GE Healthcare) at 200 V const. in 1 x TBE and collected every hour. Before lyophilization, the RNA was dialyzed against 1 M NaCl, o.n., 4 °C and twice against DEPC H_2O, o.n., 4 °C.

2.10.6 Isotopic labeling of RNA

RNAs for EMSA-binding studies were radioactively labeled. In case of *in vitro* transcribed RNAs, 5'-end phosphates of 13 pmol RNA were removed by using 2 U FastAP™ thermosensitive alkaline phosphatase (Thermo Fisher). 20 µl final reaction volume, including 20 U RNase inhibitor (SUPERase® In™, Thermo Fisher) in 1 x Tango buffer with BSA (Thermo Fisher) were incubated at 37 °C for 30 min. Dephosporylated RNA was phenol-chloroform-isoamylalcohol extracted and precipitated by adding 0.1 V 3 M NaOAc, 3.5 V EtOH and subsequent chilling at -20 °C for ≥ 30 min.

10 pmol of either chemically synthesized or *in vitro* transcribed RNA oligonucleotides were phosphorylated at the 5'-end with γ-^{32}P ATP (Hartmann Analytic) by T4 polynucleotide kinase (New England Biolabs) with 1 x buffer A in a final volume of 20 µl. Labeling reaction was carried out at 37 °C and stopped after 30 min by incubation at 75 °C for 10 min. Radiolabeled RNA was separated from free nucleotides by using a NucAway™ Spin column (Ambion) according to the manufacturer's manual. After denaturation for 10 min at 75 °C and refolding via snap cooling on ice, labeled RNA was diluted to a final concentration of 100 nM in DEPC H$_2$O and stored at -20 °C.

2.10.7 Electrophoretic mobility shift assay

Protein-RNA complexes were formed in RNase-free buffer containing 20 mM Hepes, pH 7.8, 200 mM NaCl, 4 % (v/v) glycerol, 2 mM MgCl$_2$, 2 mM DTT. Protein(s) at the indicated final concentrations, supplemented with 30 µg/ml yeast tRNA as competitor and 5 nM γ-^{32}P-labelled RNA were incubated in a final volume of 20 µl for 30 min at RT. 17 µl of the reaction mixtures were loaded on 4 % native TBE-PAGE gels and the protein-RNA complexes resolved in 1 x TBE running buffer at 70 V for 61 min. Gels were fixed for 10 min in 30 % (v/v) methanol, 10 % (v/v) acetic acid, subsequently dried in a gel dryer (BioRad) and analyzed with radiograph films (Kodak) in a Protec Optimax developer (Hohmann). In case of quantification, EMSAs were exposed to a phosphor imaging plate for ≥ 1 h and scanned with a FLA-3000 (FUJIFILM) system and analyzed using ImageJ [182] version 1.50i. Each EMSA was at least performed three times.

2.10.8 *In vitro* pull-down assay

Protein samples were mixed in their correct stoichiometric ratios, using 10 mM She2p wt/ ΔhE, 10 mM She3p-His$_6$ and 5 mM GST-Myo4-C in a final volume of 100 ml pull-down buffer containing

20 mM Hepes pH 7.8, 140 mM or 200 mM NaCl, 2 mM MgCl$_2$ and 2 mM DTT. After centrifugation for 10 min, 16,100 x g, 4 °C, 95 µl of the supernatant were incubated with 45 µl Glutathione Sepharose beads (GE Healthcare) for 30 min at 4 °C on an overhead shaker. Binding reactions were washed four times with 200 µl pull-down buffer and each time spun down for 1 min at 400 x g, 4 °C. The last washing step was performed with 41 µl pull-down buffer. Bound proteins were eluted with 41 ml pull-down buffer, supplemented with 10 mM glutathione (reduced). In pull-down experiments at room temperature and at 30 °C all experimental steps were performed at the indicated elevated temperature. 10 % of the input, 20 % of the last wash step, and 20 % of the elution were analyzed on SDS-PAGE gels (section 2.10.3).

2.10.9 Analytical size-exclusion chromatography

Analytical SEC was used to screen different combinations of She2p variants, She3p fragments and *ASH1*-E3 RNA species on their ability to form stable complexes. Unless stated otherwise, 20 µM of She2p, 20 µM She3p and 10 µM RNA were mixed in SEC buffer containing 20 mM Hepes pH 7.8 and 200 mM NaCl. For testing She2p, She3p-His$_6$ and Myo4-C on stable complex formation, the respective proteins were mixed at concentrations of 40 µM, 40 µM and 20 µM, respectively in 20 mM Hepes pH 7.8, 200 mM NaCl, 2 mM MgCl$_2$, 2 mM DTT. Components were incubated for 10 min at room temperature and spun down 10 min, 16,100 x g, 4 °C. A total volume of 200 µl was subjected to a pre-equilibrated Superose 6 (10/300) GL column (GE Healthcare) with a flow-rate of 0.5 ml/min on an ÄKTApurifier 10 (GE Healthcare) and 0.5 ml fractions were collected. The protein content of 2 % (v/v) of each eluted fraction was analyzed by SDS-PAGE, the nucleic acid content by denaturing PAGE (see section 2.10.3).

2.10.10   Circular Dichroism (CD) Spectroscopy

CD spectra for the comparison of She2p wild type and She2p E172A, F176A secondary structure content were recorded on a Jasco J-715 spectropolarimeter (JASCO) in the range of 190 nm to 260 nm at 20 °C, using a high precision quartz cuvette (Hellma Analytics) with a path length of 1 mm. Proteins were present at a concentration of 7.5 µM in a 50 mM sodium phosphate buffer pH 7.8 containing 200 mM NaCl and 1.5 mM TCEP. Measurement parameters included a scanning speed of 50 nm/min and 5 scans. The response time was 8 s. The molar ellipticity i.e. the one in relation to the

average molecular weight of the amino acids of the protein Θ_{MRW} was calculated with the formula $\Theta_{MRW} = \frac{\Theta \cdot M \cdot 100}{c \cdot d \cdot N_{aa}}$, where Θ is the measured ellipticity in [mdeg], M the molecular weight of the protein in [kDa], c the protein concentration in [mg/ml], d the thickness of the cuvette in [cm] and N_{aa} the number of amino acids of the protein.

2.10.11 Limited proteolysis

A sample of 25 nmol She2p (6-240 C-S), 30 nmol She3p(334-425)His$_6$ and 12.5 nmol *ASH1*-E3 (51 nt) in a total volume of 400 µl SEC buffer (20 mM Hepes pH 7.8, 200 mM NaCl) was subjected to size-exclusion chromatograpy (see 2.10.9) and the ternary complex containing fractions concentrated using an Amicon®Ultra centrifugal filter (30,000 Da cut-off, Merck Millipore). Several complex dilutions were analyzed on SDS-PAGE and an amount of 2 µl pure complex determined to be sufficient for proper band visualization on gels. Depending on the amount of protein complex being digested, 3 µl or 9.6 µl of 0.1 mg/ml, 0.01 mg/ml or 0.001 mg/ml trypsin, subtilisin, Glu-C protease, chymotrypsin and elastase dilutions in protease dilution buffer (20 mM Hepes pH 7.5, 50 mM NaCl, 10 mM MgSO$_4$) were added to 2 µl or 6.4 µl protein-RNA complex in a total volume of 10 µl and incubated on ice. The reaction was stopped after 30 min by the addition of 4 x SDS-loading dye. Samples were analyzed on SDS-PAGE gels. Promising bands were excised and the proteins eluted from the gel with 1 % (v/v) methanol on a pre-equilibrated PVDF membrane (Roth). After 2 d of incubation at room temperature the membrane was five times washed with each 10 % (v/v) methanol and ddH$_2$O. Dr. Reinhard Mentele (Max-Planck Institute of Biochemistry, Martinsried, Germany) performed Edman degradation [183] and sequencing.

2.10.12 Static light-scattering

SLS experiments of She2p, She3p and Myo4-C complexes were performed after size-exclusion chromatography with a Superose 6 10/300 GL column (GE Healthcare) and a flow rate of 0.1 ml/min at 4 °C, using a 270 Dual Detector and a VE3580 RI Detector (Malvern). System calibration was performed with 100 µl BSA at a concentration of 4 mg/ml. Sample concentrations were in the range of 1.9 to 3.1 mg/ml in a total volume of 100 µl. Complexes were assembled at stoichiometric ratios and applied to size-exclusion chromatography. For data analysis OmniSEC 5.02 software (Malvern)

was used. The average value of two independent experiments was used for molecular weight determination.

2.11 Bioinformatics tools

DNA and protein sequences were obtained from the *Saccharomyces* genome database [184]. Secondary structure predictions of RNAs were done with the RNA folding form of the Mfold web server [144] and secondary structures visualized with the Java applet VARNA (version 3.9, http://www.varna.lri.fr). JPred 3 [185] and PSIPRED v3.3 [186] predicted secondary structures of proteins. Physico-chemical properties of proteins such as the theoretical isoelectric point, molecular weight and the extinction coefficient were computed using the ProtParam tool of the ExPASy server [180]. Multiple sequence alignments were performed with ClustalW [187].

2.12 Structural biology

2.12.1 Small angle X-ray scattering

SAXS measurements with *ASH1* E3 (28 nt-loop) RNA from large scale *in vitro* transcriptions (see section 2.10.5) were done together with Ralf Stehle (Department of Chemistry, TUM, Munich, Germany) on a Rigaku BIOSAXS1000 instrument with a HF007 microfocus generator equipped with a Cu-target at 40 kV and 30 mA. Transmissions were measured with a photodiode beamstop, q-calibration was made by a silver-behenate measurement. Measurements were done in eight 900 s frames, tested for beam damage and averaged as images. Circular averaging and background subtraction was done with the Rigaku SAXSLab software v 3.0.1r1., ATSAS package (v 2.7.0-1 [188]). Theoretical scattering curves for RNA models were back-calculated with the program Crysol, ATSAS package (v 2.7.0-1 [188]). Plots were generated with Origin 9.1 (OriginLab). RNA was measured with 1.5, 7, 11 and 20 mg/ml in 20 mM Hepes pH 7.8, 200 mM NaCl, 2 mM $MgCl_2$ at 5 °C.

2.12.2 Nuclear magnetic resonance

NMR measurements were performed by Dr. Andreas Schlundt (Center for Integrated Protein Science Munich at the chair of Biomolecular NMR Spectroscopy, Department of Chemistry, TUM, Munich,

Germany). Therefore *in vitro* transcribed E3 (28 nt-loop) and (42 nt-TL/TLR) RNA samples (section 2.10.5) were prepared by dissolving the lyophilized RNA in 20 mM Hepes, 200 mM NaCl, 2 mM MgCl$_2$, pH 7.8 in DEPC H$_2$O with 10 % D$_2$O. The solutions were snap-cooled by boiling at 95 °C for 5 min and transferred to an ice-cold bath for 10 min before use to avoid thermodynamically favored duplex formation. E3 (28 nt-loop) and (42 nt-TL/TLR) RNAs were probed for secondary structure-analyzing imino-^{1}H-1D NMR spectra as described previously [189]. The RNA-NMR data were measured at different temperatures and imino resonances were assigned through imino-NOESY experiments recorded at 278 K. All experiments were performed using a Bruker® AVIII spectrometer at 900 MHz proton Larmor frequency equipped with a triple-resonance-cryogenic probe. Sample concentrations of E3 (28 nt-loop) and (42 nt-TL/TLR) RNA were 172 and 200 µM, respectively. Imino-NOESY experiments were recorded with 300 ms mixing time using 2048 x 220 points and 400-800 scans. Data were acquired and processed with the program TopSpin 3.5 (Bruker) and further analyzed using the program SPARKY [190]. Imino protons were assigned manually guided by the base pairing observed in the E3 42-mer crystal structure and suggested by Mfold version 2.3 [144].

2.12.3 Preparing RNA for crystallization

Unless stated otherwise, PAGE-purified chemically synthesized RNA (Dharmacon) was deprotected, desalted and lyophilized (see section 2.8.6). Scaffold RNAs like E3 (42 nt-TL/TLR) RNA were solubilized in a buffer containing 10 mM sodium cacodylate pH 6.5, 5 mM MgCl$_2$, unfolded for 10 min at 70 °C and refolded via snap cooling on ice. RNA for co-complex crystallization, like E3 (28 nt-loop), was solubilized in DEPC H$_2$O after lyophilization. Integrity of the RNA was confirmed using a 10 % denaturing TBE-PAGE gel.

2.12.4 General procedure for crystallization

A general procedure, which was applied for crystallizing RNA and RNA/protein complexes, is described in the following.

First, highly pure protein (≥ 95 %) was produced (section 2.9). RNA samples were refolded and treated as described in section 2.12.3. RNA/protein complexes were assembled by mixing single components at a molar ratio of She2p:RNA = 2:1.2, or in case of She2p:She3p:RNA at a molar ratio of 2:2.3:1. The latter complex was subjected to size-exclusion chromatography (section 2.10.9) and

sample homogeneity assessed by dynamic light scattering, if necessary. Before crystallization, RNA/protein samples were spun down ≥ 10 min, 16.100 x g, 4 °C.

To determine the appropriate RNA/protein concentration for crystallization screens, either a Pre-Crystallization Test (PCT) (Hampton Research) was performed according to manufacturer's instructions or the sample concentration was adjusted after judging drops from initial plates. If more than 80 % of the drops showed heavy precipitate directly after set-up, the concentration was reduced by 50 %. In case of more than 80 % clear drops, the concentration was doubled.

To identify initial crystallization conditions, a range of initial screens (JCSG+, Nucleix, PACT, Natrix, Classics I and Classics II) was set up. 200 nl RNA/protein were mixed with 200 nl reservoir solution on 96-well MRC Crystallization Plates™ (Molecular Dimensions) in a sitting-drop approach using the crystallization robotic system Mosquito (TTP LabTech). The reservoir volume was 80 µl. In the beginning plates were set by the MPIB Crystallization Facility (Max Planck Institute of Biochemistry, Martinsried, Germany) utilizing the PHEONIX crystallization robot (Art Robbins Instruments). Plates were stored at 21 °C and 4 °C and regularly inspected using a Leica M165 C microscope. To distinguish between protein and salt crystals in case of doubt, drops were supplemented with 0.1 V izit crystal dye (Hampton research) and inspected after 1 h.

Initial hit conditions were further optimized in refinement screens. Stock solutions to reproduce and vary the initial crystallization condition were filtered using a 0.22 µm filter (Merck Millipore). The drop volume was increased to a total volume of 2 µl and in individual cases even up to 4 µl. Usually the hanging drop vapor diffusion method was applied. VDX plates (Hampton Research) were filled with 500 µl precipitant solution per each of the 24 wells. Two parameters of the initial hit condition were varied simultaneously in either row or column of the plate. In this respect a refinement matrix was generated by altering for instance pH, salt or precipitant concentration. Additionally protein concentration was lowered and/or increased.

Crystals were harvested with cryogenic loops and passed through cryogenic protectant. Therefore usually 25 % (v/v) ethylene glycol or glycerol was added to the crystallization condition. Other cryogenic protectants like low viscosity Cryo Oil (MiTeGen), MPD or glucose, the latter ranging between 20 and 30 % were tested in some cases. Crystals were immediately flash-frozen in liquid nitrogen and stored at -80 °C until being measured.

Data collection was performed either at European Synchrotron Radiation Facility (ESRF, Grenoble, France), Swiss Light Source (SLS, Villigen, Switzerland) or Deutsches Elektronen Synchrotron (DESY, Hamburg, Germany). Details on data collection and structure determination are provided in the respective results parts.

2.12.5 Additive and heavy atom screen

In case of irregularly shaped crystals, an Additive Screen™ (Hampton Research) was used to optimize the crystallization condition. 96 additive reagents, consisting of 80 non-volatile and 16 volatile samples were included 1:10 (v/v) in the total drop volume of 700 nl during crystal set-up with the Mosquito (TTP LabTech) crystallization robot. The reservoir solution set-up was done according to the manufacturer's user guide.

The so-called heavy atom screen is a lab-made screen, consisting in addition to eight heavy atom derivatives also of several amino acids and sugar derivatives that are not included in commercial additive screens. The composition of the screen can be inspected in Appendix Table 5.1. The screen set-up was performed as for non-volatile reagents from the Hampton Additive Screen™.

2.12.6 Seeding

Seeding was either applied when single, promising crystals grew and spontaneous nucleation seemed to be a problem, or to improve small multi-crystals to grow larger single crystals.

In case of streak seeding, the crystal was touched with the Seeding Tool (Hampton Research) and sequentially passed through five drops. Those drops were set up 24 h before seeding by mixing 1 µl RNA/protein complex and 1 µl reservoir solution, using the hanging-drop vapor diffusion method [191] in 24-well VDX plates (Hampton Research) with a reservoir volume of 500 µl.

In case of micro seeding, a crystal seed stock was prepared with the Seed Bead Kit™ (Hampton research). 50 µl of well solution and a drop containing crystals was applied to the micro centrifuge tube containing a PTFE bead. After vortexing for 1.5 – 3 min either streak seeding was performed, or 450 µl of well solution were added to produce the seed stock. Drops were set up with serial seed dilutions (1:10, 1:100, 1:1000) by mixing 1 µl RNA/protein complex and 1 µl reservoir solution.

2.12.7 *In situ* proteolysis

The complex of She2p (6-246, C-S), She3p-His$_6$ and *ASH1*-E3 (28 nt-loop) was mixed at a molar ratio of 2:2.3:1 in 20 mM Hepes/NaOH pH 7.8, 200 mM NaCl, 2 mM DTT in a total volume of 400 μl. The stoichiometric complex was separated on a Superose 6 10/300 GL column (GE Healthcare) in the same buffer. She2p-She3p protein content was verified on a 12 % SDS-PAGE gel, whereas RNA was analyzed on a 10 % urea PAGE gel (section 2.10.3). For *in situ* proteolysis [192] 1 mg/ml α-chymotrypsin and 1.5 mg/ml trypsin stock solutions were prepared in 1 mM HCl and 2 mM CaCl$_2$. Either of the protease was added to 2.11 mg/ml She2p-She3p-RNA complex at a concentration of 1:1000 (w/w) or 1:5000 (w/w) and initial crystallization trials were set at RT and 4 °C as described in section 2.12.4. As soon as crystals reached a certain size that made their harvesting possible, they were immediately cryogenically protected and flash-frozen at -80 °C.

2.12.8 Dehydrating crystals

Native crystals of She2p (6-246, C-S) and E3 (28 nt-loop) (see section 3.3.4.2) were used for dehydration [193]. Each drop containing She2p/RNA crystals was equilibrated for 24 h against a reservoir with 500 μl well solution and consecutively transferred to reservoirs with 2.5 % increased precipitant concentration at 4 °C. Hence, starting from 20 %, the final crystallization solution contained 30 % (w/v) PEG3350 as dehydrating agent. Crystals from different dehydration stages were harvested, cryogenically protected and flash-frozen as described in section 2.12.4.

2.12.9 Chemical cross-linking

According to the technique described in reference [194], vapor diffusion was used to gently apply the cross-linking reagent to the She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405):E3 (28 nt-loop) crystals in the crystallization drop. Therefore 2-5 μl of either 16 % (v/v) formaldehyde or 25 % (v/v) glutaraldehyde were applied to the sample cavity of a well in a 24-well sitting drop plate (MiTeGen). The surrounding reservoir was filled with mother liquor used for crystallization and the well sealed with a coverslip containing the crystals in the crystallization drop.

In an additional set-up the above-mentioned cross-linking reagents were applied more harshly to the crystals by directly pipetting 0.2 μl of either of the chemicals to the crystallization droplet.

Cross-linking times were varied in both methods (10 min, 30 min, 1 h and 6.5 h) at either room temperature or 4 °C. Crystals from different experimental set-ups were harvested, cryogenically protected and flash-frozen as described in section 2.12.4.

2.12.10   Capillary crystallization

When testing crystal growth in capillaries, several Special Glass 10 Capillaries (Hampton Research) with varying outside diameters (0.2, 0.4, 0.6 and 0.8 mm) were used. 1 µl of each ingredient was aspirated as follows: low viscosity Cryo Oil (MiTeGen), air, crystallization solution, protein/RNA-containing sample, air, cryo oil and air. Capillaries were set in triplicates, sealed with Beeswax (Hampton Research) and stored at 4 °C.

2.12.11   Soaking

Native She2p:RNA crystals were either soaked with heavy metals, or commercially synthesized, HPLC-purified She3p peptides (JPT Peptide Technologies). In each case a final concentration of 1 mM soaking compound was applied to the drop and equilibrated for 24 h before freezing (section 2.12.4.). Beforehand, She3p peptides were solubilized in water and dialyzed against crystallization condition over night at 4 °C, using a dialysis button (Hampton Research) and a 1,000 Da cut-off ZelluTrans membrane (Roth).

2.12.12   Structure visualization and analysis

Images of the crystal structures, superimpositions of the co-complex and apo-structure, as well as electrostatic surface potentials were prepared with PyMol (version 1.7; Schrodinger LLC.; http://www.pymol.org/). Electron density visualization was done in CueMol2 (version 2.0; BKR Laboratory; http://www.cuemol.org/). B factors (Appendix Table 5.2 and Appendix Table 5.3) were calculated with the baverage program from the CCP4 suite [195]. RMSD calculations were performed using the protein structure comparison service PDBeFold (version 2.59; http://www.ebi.ac.uk/msd-srv/ssm; EMBL-EBI). Schematic diagrams of protein-RNA interactions were generated using NUCPLOT [196] from the PDBsum database (http://www.biochem.ucl.ac.uk/bsm/pdbsum). Structural models and data sets are available on the protein data bank (http://www.rcsb.org) with PDB-IDs as follows:

5M0H for the "Crystal structure of the central flexible region of *ASH1*-mRNA E3-localization element", 5M0I for the "Crystal structure of the nuclear complex with She2p and the *ASH1*-mRNA E3-localization element" and 5M0J for the "Crystal structure of the cytoplasmic complex of She2p, She3p, and the *ASH1*-mRNA E3-localization element".

3 Results

3.1 SHE complex assembly and its motility

Single-particle motility assays were planned to test if the *in vitro* reconstituted SHE complexes [145] are also functional by means of Myo4p-dependent cargo-RNA transport and to test which parameters potentially control and influence particle assembly and its motility.

3.1.1 Experimental set-up

Tested complexes for single-particle motility assays were composed of the yeast motor complex comprising Myo4p, the myosin-light chain Mlc1p, calmodulin Cmd1p and the myosin-specific chaperone She4p. Further components of the cytoplasmic SHE core complex included She2p, She3p and, depending on the experiment, up to two (different) localizing RNA elements (Figure 3.1). Single-particle motility experiments, which were carried out by Dr. Dennis Zimmermann (Department of Molecular Genetics & Cell Biology, University of Chicago, USA), should yield information about particle motility, its velocity and run length as well as possible oligomerization states (Figure 3.1 A). Particle motion in general was followed by total internal reflection fluorescence microscopy (TIRFM) by either tracing the GFP-signal, N-terminally fused to the She2p protein, or via fluorescently labeled DNA primers hybridized with the RNA (Figure 3.1 B-G). Dr. Roland Heym (formerly Niessing lab) established the initial set-up, including plasmid cloning, protein expression and purification. After his departure from the lab, I continued to optimize and modified this set-up for TIRFM.

According to the results from Heym and colleagues [145] the assembly of the SHE core complex induces Myo4p dimerization and therefore reconstituted particles should be able to fulfill active cargo-RNA transport. To test this interpretation, a fragment of the *ASH1* mRNA containing one E3 localization element (*ASH1*-3') should be assayed with GFP-She2p (Figure 3.1 B).

To clarify if the RNA-binding protein She2p or its LE RNA is responsible for complex assembly, the second set-up included GFP-She2p wild type without adding an additional RNA (Figure 3.1 C).

Since residual RNA contaminations in the supposedly RNA-free protein preparations could lead to wrong interpretations, particles bearing the RNA-binding deficient GFP-She2p mutant (N36S, R63K) [148] were also assessed (Figure 3.1 D).
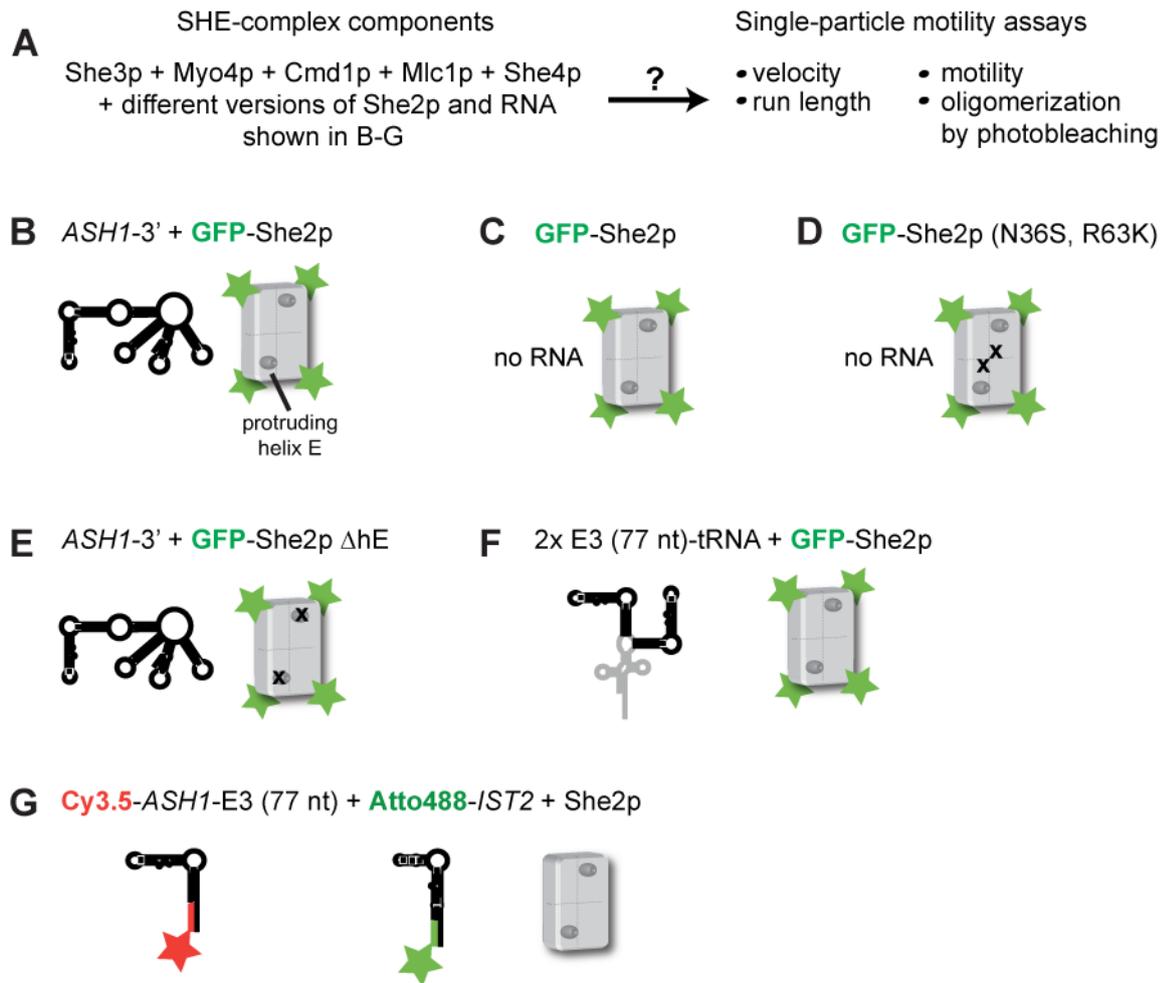
Figure 3.1: Schematic experimental set-up for single-particle motility assays. A) Particles had to be assembled in the presence of the Myo4p motor and She3p-His$_6$. Calmodulin, myosin-light chain and the myosin chaperone She4p were separately added to each experiment. Set-ups differed in RNA and She2p composition B)-G). **B)** One She2p tetramer (grey) [152] carries four N-terminal GFP tags (green stars, see also C), D), E) and F)) and was tested with *ASH1* 3' comprising one E3 localization element. **C)** GFP-She2p and **D)** the RNA-binding defective mutant GFP-She2p (N36S, R63K) were analyzed without RNA. Black crosses indicate respective mutations. **E)** GFP-She2p, lacking the protruding helix E (ΔhE, indicated by black crosses) and therefore being binding-deficient for She3p, was tested with *ASH1*-3' RNA **F)** An RNA construct bearing two LEs 2x E3 (77 nt)-tRNA was combined with GFP-She2p. **G)** Cy3.5-*ASH1*-E3 (77 nt) and ATTO488-*IST2* were used to test the simultaneous transport of two LEs from different RNAs together with She2p.

When investigating the assembly-mediating features, also the She2p-She3p interaction could in principle be essential for particle motility. This assumption could be challenged using a GFP-She2p mutant that is lacking protruding helices E (GFP-She2p ΔhE) and is therefore unable to bind She3p [139] (Figure 3.1 E).

Since the SHE complex can bind two LEs, one RNA bearing two LEs can oligomerize SHE core complexes into larger particles. This property has already been shown *in vitro* in our lab by dynamic light scattering and sucrose density gradient centrifugation experiments [145]. To see whether this clustering event also occurs in the fully assembled, motor machinery-containing particle, the same set-up was tested in single-particle-motility assays by using GFP-She2p and an RNA with two E3 LEs connected by a three-way junction [175] (Figure 3.1 F). Photo-bleaching experiments, in which the number of GFP molecules in a complex can be measured by counting the number of fluorescence-recovery steps, were performed to clarify the oligomerization state of SHE particles. An increased number of SHE complexes in a moving particle would also increase the number of incorporated Myo4 motors. This could potentially have an effect on motility and should be clarified by measuring particle run-lengths and velocities (Figure 3.1 F).

The last experimental condition comprised unlabeled She2p with two LEs from different RNAs, *ASH1*-E3 and *IST2*. By combining different RNA fluorescence labels for each of them, Cy3.5 for *ASH1*-E3 and ATTO488 for *IST2*, a simultaneous transport would be visualized by detecting superposed fluorescence signals in TRIFM experiments (Figure 3.1 G).

3.1.2   Production of different RNA constructs for SHE complex assembly

All RNAs for single-particle motility assays were produced by small-scale *in vitro* transcription (section 2.10.4). In each case the starting material were PCR fragments with a T7 RNA polymerase promoter region located upstream of the respective DNA template sequence. The first RNA construct was a fragment of the *ASH1* mRNA containing the whole E3 LE (*ASH1*-3'; Figure 3.2 A). Several bases up- or downstream of the LE were added to fluorescently label the RNA by DNA hybridization (section 2.5). Secondary-structure predictions were used to confirm that the secondary structure of the LE itself was not affected by these heterologous sequences. It turned out that the RNA, hybridized with up to three fluorescently tagged DNA oligonucleotides, was not well detectable in TIRFM experiments. Therefore it was decided to visualize the particle via GFP-labeling of She2p instead.

The second RNA construct contained two *ASH1*-E3 elements of 77 nt length [139], which were fused to each other via a three-way junction from the Twort intron that was fused to a tRNA-scaffolding tag [175]. A 24 nt single-stranded 3' overhang was added to keep the possibility to fluorescently label the RNA, but was not used afterwards (2x E3-(77 nt)-tRNA; Figure 3.2 A).

In another set-up the simultaneous transport of two different RNAs should be visualized. Therefore, the *ASH1*-E3 (77 nt)-RNA, as well as the LE of the *IST2* RNA was extended at their 3' ends to allow for hybridization of DNA oligonucleotides carrying divergent fluorescent labels, Cy3.5 and ATTO488, respectively. Quality control of the *in vitro* transcribed RNAs by assessing their denatured and native state in EMSAs proved RNA's integrity (Figure 3.2 B, C) and homogenous folding (Figure 3.2 B).
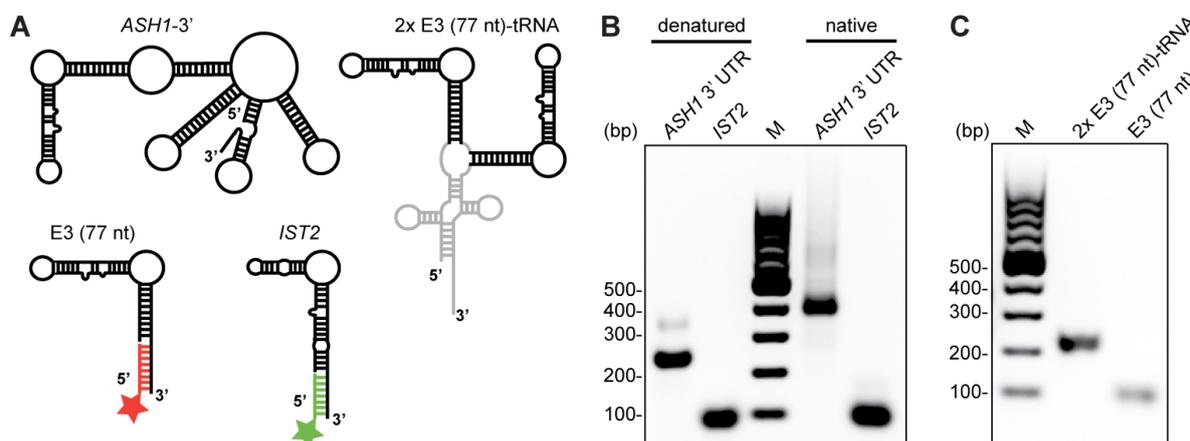
Figure 3.2: RNA constructs for single particle motility experiments and their quality control. A) Schemes of secondary structure predictions of *ASH1*-E3 (77 nt), the localizing stem loop of *IST2*, the 2x E3 (77 nt)-tRNA construct and *ASH1* 3' UTR, comprising one E3 localization element. The inserted tRNA branch in the 2x E3 (77 nt)-tRNA construct is shown in gray. If RNAs were monitored in single particle motility experiments, their single-stranded 3' ends were hybridized as indicated with either a Cy3.5-DNA oligonucleotide (red) or an ATTO488-DNA primer (green). **B)-C)** The quality of *in vitro* transcribed RNAs was assessed on agarose gels. M indicates the 100 bp DNA marker. **B)** *ASH1* 3' UTR and *IST2* RNA as denatured and native samples. One homogenously folded main species can be detected in their native state. **C)** Denatured samples of 2x E3 (77 nt)-tRNA and *ASH1*-E3 (77 nt) proof RNA's integrity.

3.1.3   Expression and purification of proteins for SHE complex assembly

Yeast calmodulin (Cmd1p) and myosin-light chain (Mlc1p) were expressed in bacteria and purified as described in section 2.9.12 (Figure 3.3 A, B). FLAG-tagged, full-length yeast Myo4p was purified as multi-protein complex together with She3p, She4p, Cmd1p, and Mlc1p, which were coexpressed in insect cells. Since calmodulin and the myosin-light chain are critical for motor protein function [197], they were also added during Myo4p purification (section 2.9.10). The concentration of the multi-protein complex was estimated by comparing band intensities of different Myo4p dilutions with known amounts of BSA in SDS-gels (Figure 3.3 C). She2p wild type and the RNA-binding deficient mutant She2p (N36S, R63K) were expressed with a GST-tag in bacteria and purified as essentially described in section 2.9.6. An initial affinity-capture step, followed by GST-tag-cleavage, anion exchange and size-exclusion polishing yielded purities of ≥ 95 % (Figure 3.3 D, E).

The necessity of cargo-RNA for complex assembly can just be assessed reliably if protein components themselves are completely RNA-free after purification. Although the normal production procedure for full-length She3p-His$_6$ comprises a high-salt washing step to remove the RNA, it is also coexpressed with She2p wild type in baculovirus-infected insect cells. Possible She2p remnants in the She3p preparation would still be able to bind RNA and as consequence interfere with supposedly RNA-free experimental set-ups (Figure 3.1 C, D). To overcome this problem, two insect cell vectors were generated pFBDual_*SHE3-His6_SHE2_N36S_R63K* and pFBDual_*SHE3-His6_SHE2_ΔhE* (section 2.3, Table 2.3), both coexpressing She3p full-length and either an RNA- or She3p-binding deficient She2p mutant. Viruses containing the first construct had very low infection rates and therefore little protein yields. When infecting High Five cells with viruses containing the second construct, cells were dying. Therefore a third approach was used, where the purification of full-length She3p-His$_6$ was extended by an additional competition wash step with She2p (N36S, R63K) to completely remove She2p wild-type remnants as described in section 2.9.7 (Figure 3.3 F).

Based on pET28a_*eGFP_SHE2*, pET28a plasmids coding for His$_6$-GFP-She2p (N36S, R63K) and His$_6$-GFP-She2p (ΔhE) were cloned (section 2.3, Table 2.3). Proteins, expressed in bacteria were purified as described in section 2.9.6. In brief, His$_6$-GFP-She2p variants were affinity-captured by a HisTrap column, further purified by anion exchange and polished by size-exclusion chromatography. After the final purification step, purest fractions based on inspection of SDS-PAGE gels were pooled and concentrated (Figure 3.3 G-I). For She2p and She3p, the successful removal of contaminating nucleic acids was confirmed by measuring their absorbance at 260/ 280 nm (section 2.10.2) after purification.
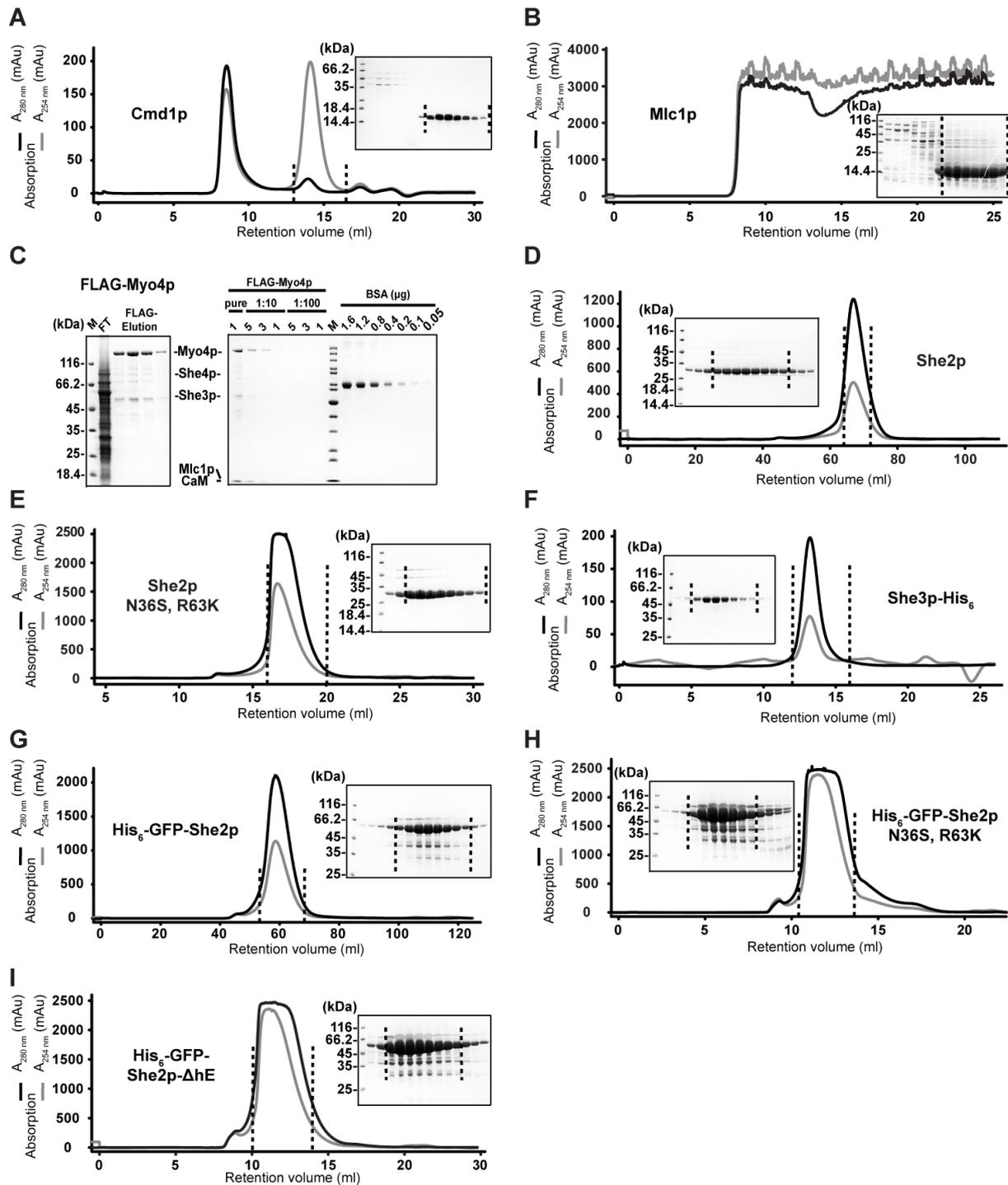
Figure 3.3: Protein purifications of SHE-complex components for motility experiments. A), B) and **D)-I)** Size-exclusion chromatograms of the final protein purification steps with their according SDS-PAGE gels. Purest fractions, delimited with dashed lines, were pooled and concentrated. **A)** shows Cmd1p (16.3 kDa) on Superose 12 10/300 GL, **B)** Mlc1p (16.6 kDa) on Superdex 200 16/60 GL. **C)** SDS-PAGE gel of FLAG-Myo4p (170.4 kDa) purification from insect cell extract. FT indicates the flow-through sample of the FLAG-bead capture step, showing efficient target protein binding. She4p (89.5 kDa), She3p-His$_6$ (48.2 kDa), Mlc1p (16.6 kDa) and Cmd1p (16.3 kDa) were co-purified (left). Concentration estimation of the multi-protein sample was done by comparing band intensities of different protein dilutions with known

amounts of BSA (right). Triplicate determination yields $c_{FLAG-Myo4p}$ = 268 ng/µl. Final purification step of **D)** She2p (28.4 kDa) and **E)** She2p N36S, R63K (28.4 kDa) on Superdex 200 16/60 GL. **F)** shows She3p-His$_6$ (48.2 kDa) on Superose 6 10/300 GL. **G)** His$_6$-GFP-She2p (60.0 kDa), **H)** His$_6$-GFP-She2p N36S, R63K (59.8 kDa) and **I)** His$_6$-GFP-She2p-ΔhE (58.9 kDa) on Superdex 200 10/300 GL.

3.1.4   Reconstituted SHE complexes move processively along actin *in vitro*

Results of the following single-particle motility assays derive from Dr. Dennis Zimmermann (Department of Molecular Genetics & Cell Biology, University of Chicago, USA), who also prepared fluorescent tetramethylrhodamine (TMR)-labeled actin filaments. TIRFM experiments with *in vitro* reconstituted particles were set up as described in section 0 and their movement was traced by following the green GFP signal on the red TMR-actin cables. In line with the reconstitution experiments described in reference [145] that showed an incorporation of two Myo4p motors in one SHE particle, probing of GFP-She2p together with *ASH1*-3' RNA indeed resulted in particles that were actively transported along actin filaments. The mean velocity of inspected particles was 0.74 ± 0.20 µm/s (mean ± standard deviation) with a run length of 2.13 ± 0.89 µm (Appendix Figure 5.2 A, E and F). Respective movies for all motility assays can be retrieved from the online supplemental material of reference [145].

To assess the amount of She2p molecules per motile particle, a subset of particles was subjected to photobleaching experiments. As an internal control, the bleaching-step profiles of GFP-She2p alone were analyzed (Appendix Figure 5.3 A, C). Since She2p is active as a tetramer, four bleaching steps should be observed for each functional She2p oligomer. In the experiments, between one and four bleaching steps could be observed. An explanation is given by the fact that She2p alone exists in equilibrium between dimeric and tetrameric states [150], [198]. Additional counts of one and three molecules result from the fact that some fluorophores are usually inactive. Furthermore, bleaching step profiles of moving particles with a single localization element were recorded (Appendix Figure 5.3 B, C). Roughly 4.5 times more particles containing tetrameric GFP-She2p molecules were counted in the fully assembled motile complex than in particles with GFP-She2p tetramers alone (Appendix Figure 5.3 C). This observation was in accordance to previous results showing that She2p becomes stabilized in the fully assembled, motile complex [139]. Since not more than four bleaching steps were observed in any of the inspected particles (Appendix Figure 5.3 C), it can be concluded that particles with a single localization element indeed consist of a single SHE complex.

3.1.5 RNA cargo is dispensable for processive particle transport

To answer the question whether cargo RNA is required for processive movement, particles containing GFP-She2p, but lacking the RNA were assembled and traced in motility assays by following the GFP signal. The mean run length was determined to be 2.48 ± 1.02 µm, and the mean velocity of moving particles was 0.65 ± 0.15 µm/s (Appendix Figure 5.2 B, E and F). These values are in the same range as the velocity and run length of particles in the presence of RNA (Appendix Figure 5.2 A, E and F), suggesting that cargo RNA itself does not activate motility and thus complex motility.

This interpretation is limited by the fact that even traces of remaining RNA contaminations in the supposedly RNA-free protein preparations could alter the results. Residual RNAs could bind to the She2p:She3p complex and potentially activate motility. To tackle this problem, particles bearing the RNA-binding deficient GFP-She2p mutant (N36S, R63K) [148] were assessed in motility assays without the addition of RNA (Appendix Figure 5.2 C, E and F). Observing a mean velocity of 0.62 ± 0.20 µm/s and a run length of 2.02 ± 0.69 µm, the tested particles showed comparable motility parameters as the wild-type complexes. Thus, we can conclude that RNA *per se* is dispensable for complex assembly and the processive movement of SHE particles.

3.1.6 She2p:She3p interaction is essential for motility

When investigating the assembly-mediating features, in principle the She2p-She3p interaction could also be essential for particle motility. This assumption was challenged using particles containing the She3p-binding deficient mutant GFP-She2p ΔhE [139] and *ASH1*-3' RNA (Appendix Figure 5.2 F). In this set-up the assembled particles were not able to move along actin filaments, proving that the protein-protein interaction between She2p and She3p is essential for SHE-complex motility.

3.1.7 SHE complexes with a two-zip code-RNA show no increased processivity

Particles containing GFP-She2p in combination with 2x E3-(77 nt)-tRNA were assessed in TIRFM experiments to understand the effect of an RNA containing two localization elements on the particle oligomerization state and its motility parameters. Moving particles containing the double-LE RNA appeared on average brighter than those with just one LE. This observation is substantiated by results from photo-bleaching experiments, showing that over 65 % of the analyzed 2x E3-(77 nt)-tRNA particles had more bleaching steps than the majority of particles with a single LE RNA. In line

with this, particles with such two-LE containing RNAs resulted in twice as many fluorescence-intensity populations than particles assembled with *ASH1*-3' (Appendix Figure 5.3 D, E). Taken these results together, the oligomerization of SHE-core complexes by two identical LEs in one RNA as shown by *in vitro* reconstitution experiments [145], also occurs in the fully assembled, motor machinery-containing particle. Keeping in mind that the *ASH1* mRNA has four LEs, even larger particles should be expected *in vivo*.

Although one could assume that an incorporation of more than two Myo4p motors would influence particle processivity, the mean velocity (0.54 ± 0.13 µm/s) and run length (1.93 ± 0.68 µm) of multimeric particles (Appendix Figure 5.2 D-F) was in the same range as for particles with a single SHE complex. These results indicate that particle oligomerization has no major influence on its motility parameters.

3.1.8 SHE complexes can simultaneously transport two RNA species

To test whether the reconstituted particles are also able to simultaneously transport two different RNAs, as already shown by live-cell imaging experiments [199], particles were assembled in the presence of She2p, Cy3.5-labeled *ASH1*-E3-77 and ATTO488-labeled *IST2*-LE RNA. TIRFM experiments showed that besides red and green particles, carrying either *ASH1*-E3-77 or *IST2*, also some yellow particles, containing the two of them, moved actively along the actin filaments. Although the number of motile particles with both RNAs bound in one complex was too low for robust quantitative statistical evaluation (Appendix Figure 5.2 F), this experiment qualitatively proofs that (i) two different RNAs can be transported in one reconstituted particle and (ii) that the cargo RNA is definitely incorporated in the reconstituted complex.

3.2 Core complex stability in absence of RNA

Motility assays showed that RNA is dispensable for complex assembly and the processive movement of SHE particles (section 3.1.5). This observation was contradictory to a previous study [159]. While Sladewski and colleagues claimed that the mRNA might be essential for motility, our results show the opposite. Moreover the authors stated in a recent commentary that particle movement in absence of RNA must have been artificially induced by protein-protein interactions at low ionic-strength (50 mM KCl) conditions and that "A stable Myo4p-She3p-She2p complex does not form at 140 mM KCl *in vitro*

and thus Myo4p would be non-motile in the cell without cargo." [165]. Since no references were given to support this statement Annika Niedner-Boblenz (Niessing lab) and me directly assessed the stability of the *ASH1* mRNP-core complex in absence of cargo RNA at different conditions.

3.2.1 She2p, She3p and Myo4-C form stable complexes in absence of cargo RNA in pull-down experiments

In order to understand the stability of the SHE complex in absence of RNA in more detail, a series of *in vitro* pull-down experiments were performed (Figure 3.4). Proteins were purified as described in sections 2.9.6, 2.9.7 and 2.9.11, and ensured to be RNA-free by measuring the $A_{260\ nm}/A_{280\ nm}$ ratio. Since pull-down experiments with Ni-sepharose and She3p-His$_6$ as bait showed unspecific protein binding to the beads, GST-Myo4p-C and glutathione sepharose beads were used instead (for experimental details see section 2.10.8). Complex stability of She2p, She3p and Myo4p-C was tested at three different salt concentrations. A buffer containing 50 mM sodium chloride was chosen because it was identical to the ionic strength in single-particle-motility assays. Set-ups with 140 mM sodium chloride resembled the physiologic ionic strength level in the cell. To draw a more complete picture of salt dependence on complex stability an elevated salt concentration of 200 mM sodium chloride, as has been used in other assays before [139], was also tested.

Pull-down experiments at 4 °C showed that in absence of RNA, She2p interacts with She3p-His$_6$ and Myo4p-C at a physiological level of 140 mM sodium chloride. An elevated ionic strength of 200 mM sodium chloride yielded the same result. At a salt concentration of 50 mM a small proportion of She2p and She3p-His$_6$ was still stable and co-eluted with Myo4p-C (Figure 3.4 A). In contrast, the She3p-binding deficient mutant She2p ΔhE [139] failed to form a complex with Myo4p-C and She3p-His$_6$ independent from salt concentration, showing that the She2p-She3p interaction is indeed specific (Figure 3.4 B). Also She3p-His$_6$ and Myo4p-C alone formed stable complexes at different ionic strengths (Figure 3.4 C). Consequently Myo4p-C failed to interact with She2p in absence of She3p-His$_6$, irrespective of the salt concentration (Figure 3.4 D). When performing the same pull-down experiments at room temperature no difference could be detected. Further control experiments were conducted to ensure that the observed protein interactions (Figure 3.4 A-C) are not due to artifacts caused the GST-tag on Myo4p-C (Figure 3.4 E, F). Pull-down experiments were also performed at conditions including the optimal growth-temperature of yeast (30 °C) and a salt concentration of 140 mM sodium chloride, to more closely resemble the physiologic environment

(Figure 3.4 G, H). Here, She2p also co-eluted with Myo4p-C in presence of She3p-His$_6$ and formed a stable complex in absence of RNA, whereas She2p ΔhE again failed to interact with the respective components (Figure 3.4 G). Control experiments showed that Myo4p-C bound to She3p-His$_6$, while She2p and its mutant ΔhE did not (Figure 3.4 H).

Together these results clearly demonstrate that the observed complex formation of Myo4p, She2p and She3p in the absence of RNA is stable and specific at near-physiologic conditions, as well as at elevated ionic strength levels. In agreement with single-particle motility assays, this specific interaction is also visible at a low salt concentration of 50 mM sodium chloride.
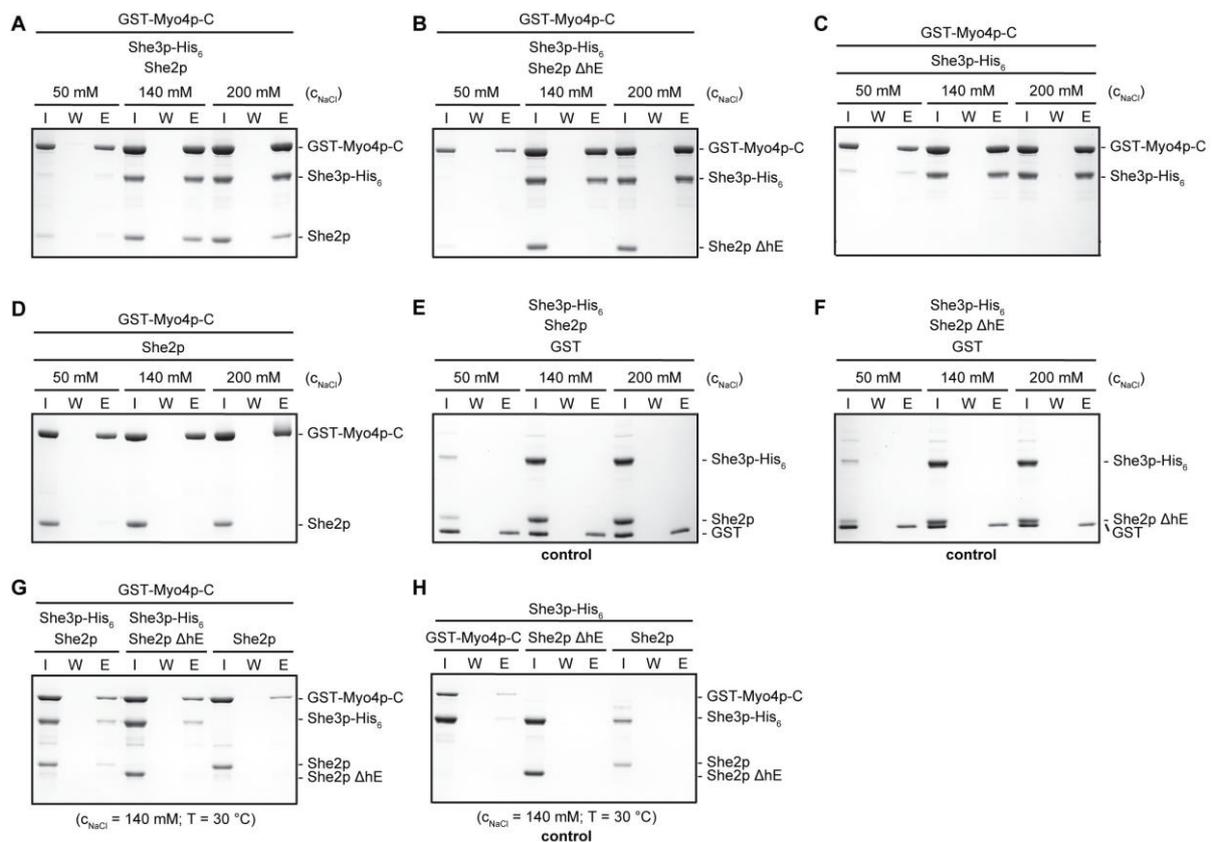
Figure 3.4: GST pull-down experiments at different salt concentrations and temperatures. Except for control experiments, GST-Myo4p-C served as bait. I = Input, W = Wash, E = Elution. **A)-F)** Pull-downs at 50 mM, 140 mM and 200 mM sodium chloride concentration and 4 °C. Protein-protein interactions were examined between GST-Myo4-C and **A)** She2p and She3p-His$_6$, **B)** She2p ΔhE mutant and She3p-His$_6$, **C)** She3p-His$_6$, **D)** She2p. **E)-F)** Control experiments that probe the bait-tag GST interaction with **E)** She3p-His$_6$ and She2p or **F)** She3p-His$_6$ and She2p ΔhE. **G)-H)** Pull-down experiments at near-physiologic conditions, resembled by 30 °C and 140 mM sodium chloride. **G)** The Interaction of GST-tagged Myo4-C was tested with She3p-His$_6$ and She2p, She3p-His$_6$ and She2p ΔhE or with She2p alone. **H)** Control experiments comprising She3p-His$_6$ in combination with either GST-Myo4p-C, or She2p ΔhE, or She2p. All pull-downs were repeated at least twice.

3.2.2 Protein-protein binding events are robust against an elevated ionic strength level

To further test the robustness of Myo4p-C, She2p and She3p-His$_6$ complex formation without RNA, analytical size exclusion-chromatography experiments (section 2.10.9) were performed (Figure 3.5 A). Consistent with the results from pull-down experiments, all of the three protein components eluted in one stable complex after about 24 minutes (12 ml) of chromatography in a buffer containing 200 mM sodium chloride. The fact that She2p ΔhE was not able to join the Myo4p-She3p co-complex again underlines the specificity of the ternary protein complex.
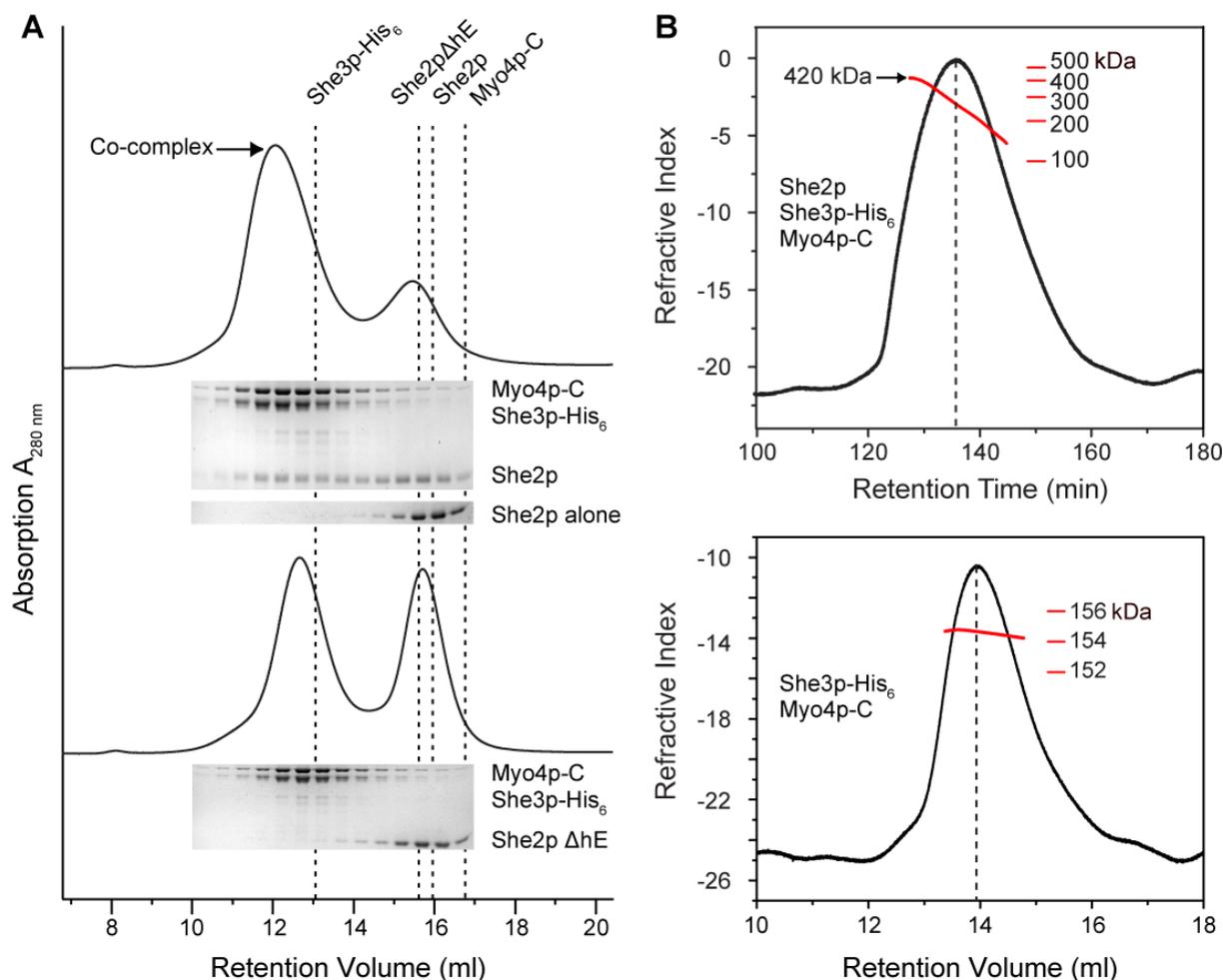
Figure 3.5: She2p-She3p-Myo4p co-complex assessment. A) Size-exclusion chromatography and SDS-PAGE from corresponding elution fractions. She2p co-elutes with She3p-His$_6$ and Myo4p-C in absence of RNA (upper part), whereas its mutant ΔhE does not (lower part). As control, fractions of the elution peak of She2p alone are shown in a separate SDS-PAGE gel (upper part). Dashed lines resemble elution peak maxima of single protein components as indicated. **B)** Static light-scattering experiments. Molecular mass distributions over complex peaks are depicted as red lines. The upper panel shows assessment of the co-complex consisting of She2p, She3p-His$_6$ and Myo4p-C. The median molecular weight of the peak was determined to be 248 kDa, while a portion of the peak content had a mass of 420 kDa. In a control experiment (lower panel), She3p-His$_6$ and Myo4p-C formed a complex of 154 kDa. Two independent experiments were used for the molecular weight determination.

Static light-scattering experiments (section 2.10.12) were used to determine the mass of complexes (Figure 3.5 B). While the median molecular weight of the She2p, She3p-His$_6$ and Myo4p-C co-complex elution peak was 248 kDa, an initial eluting portion of the peak content had a mass of 420 kDa. This mass is close to the expected molecular weight of the stoichiometric complex consisting of four She2p, four She3p and two Myo4p-C molecules counting 434 kDa. Although the mature complex disassembles over time, parts of it were still stable even after more than 120 min in high salt buffer (200 mM sodium chloride). A median molecular weight of 154 kDa was determined in a control experiment, probing She3p-His$_6$ together with Myo4p-C in size-exclusion chromatography. Since one She3p dimer is bound by one Myo4p-C monomer, the experimentally determined mass was in accordance with the expected molecular weight of the trimeric complex of 160 kDa. Taking these data together, the observed protein-protein binding events between Myo4p-C, She2p and She3p-His$_6$ in absence of RNA are robust at an elevated ionic strength level and persist even over a period of two hours.

3.3 Structural studies on the *ASH1* mRNP-core complex

The *ASH1* mRNP-core complex can be quantitatively reconstituted *in vitro* and is therefore well suited to gain a molecular understanding of how specific RNA recognition by an mRNA-transport machinery is achieved. The following sections describe the structural elucidation of a cargo RNA-localization element, its co-complex with the RBP She2p and of the highly specific ternary complex additionally bearing She3p.

3.3.1 Defining a minimal *ASH1* E3-RNA localization-element

To date the shortest fragment of *ASH1* RNA that still mediates synergistic RNA binding with She2p and She3p comprises a 51-nt long region of the E3 localization element [139]. Apart from that, Dr. Roland Heym has performed several EMSAs in order to narrow the RNA down to its essential binding fragment (Appendix Figure 5.4). A construct comprising a deletion of the upper part of the stem and of the nona-loop (nt 1786-1802) combined with an insertion of a more compact tetra-loop evolved from a combination of these previous experiments. This minimal E3 (28 nt-loop) construct was tested for synergistic RNA recognition together with She2p and full-length She3p in EMSAs (section 2.10.7, Figure 3.6 A). To ensure that just specific binding events are analyzed, a constant amount of 30 µg/ml yeast tRNA was added in each EMSA experiment. The E3-(51 nt) RNA was produced by small-scale *in vitro* transcription (section 2.10.4), while E3 (28 nt-loop) was purchased as chemically synthesized

RNA (section 2.8.6). Probing *ASH1*-E3 (51 nt) together with She2p and She3p yielded high-affinity band shifts (Figure 3.6 A, apparent K_D = 0.20 ± 0.03 µM determined in Appendix Figure 5.14 A, E). Additionally there was a second, not-shifting RNA band visible that probably constitutes a non-functional folding state. Gels with the minimal E3 (28 nt-loop) RNA showed that this construct still allows for specific recognition by the She2p-She3p complex and therefore was used in further experiments (Figure 3.6 A, apparent K_D = 0.26 ± 0.09 µM determined in Appendix Figure 5.14 B, E). In a control experiment full-length She3p-His$_6$ bound RNAs in absence of She2p just at very high She3p concentrations (Figure 3.6 B).
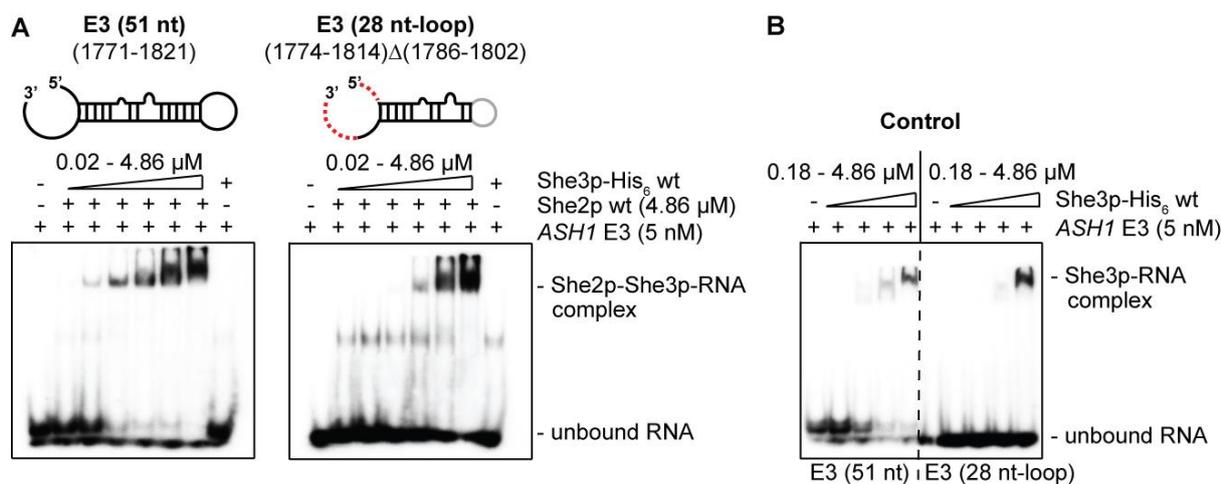
Figure 3.6: Radioactive EMSAs with *ASH1*-E3 variants. A) Radioactively labeled *ASH1* E3-(51 nt) and E3 (28 nt-loop) RNA were tested for ternary complex formation with She2p and She3p-His$_6$. RNA secondary structure predictions are shown schematically. Red, dashed lines indicate base deletions. Gray loop indicates base replacements by a GAAA insertion. She3p-His$_6$ concentrations were 20 nM, 60 nM, 180 nM, 540 nM, 1.62 µM and 4.86 µM **B)** Control EMSA showing that She3p binds RNA in absence of She2p just at very high She3p concentrations. EMSAs were performed as independent triplicates.

3.3.2 Crystallizing a localizing element

Concerning tertiary structures of LEs the only known examples of structured stem loops are the *fs(1)K10* mRNA, and SOLE in the *oskar* mRNA of *D. melanogaster*, which were elucidated by NMR experiments [73], [74]. In order to obtain a high-resolution crystal structure of a localizing element from yeast, several RNA constructs were subjected to crystallization trials. RNA sequences are summarized in Table 2.5, while the secondary structure predictions of RNA constructs used in this study can be inspected in Appendix Figure 5.5, Appendix Figure 5.6 and Appendix Figure 5.7. The general crystallization procedure is described in section 2.12.3.

3.3.2.1 Screening for crystallizable LE variants

At first, initial crystallization trials were performed with the chemically synthesized, minimal E3 (28 nt-loop) RNA. After deprotection the RNA was dissolved to 10 mg/ml in RNase-free water and tested at 21 °C in three different 96-well screens suitable for RNA crystallization (AJ1, AJ2 + Natrix and Qiagen Nucleix). RNA and the precipitant were mixed 1:1 in varying drop-sizes (400 nl and 200 nl) and the drops set by the MPIB Crystallization Facility (Martinsried, Germany). Plates were inspected regularly but yielded no crystals. While flexible parts potentially hamper crystal formation, a shortened construct E3 (25 nt-loop) lacking three flexible bases at the 3' end was assessed. After deprotection, desalting and refolding the RNA in 10 mM sodium cacodylate pH 6.5 and 5 mM $MgCl_2$, Natrix I and II initial screens were pipetted by the Mosquito dispenser (TTP Labtech). Drops containing 200 nl reservoir solution and 200 nl RNA with 6.8 mg/ml did not result in any crystals at 21 °C. To overcome this problem a general module for RNA crystallization was implemented that was shown to facilitate crystal packing by providing a platform for defined intermolecular RNA interactions [200], [201], [202]. To this end a hybrid RNA was generated in which the GAAA tetra-loop of the E3 (25 nt-loop) RNA was replaced by the heterologous 17 nt/GAAA scaffold hairpin yielding the 42 nt-long E3 tetra-loop/ tetra-loop receptor (E3 (42 nt-TL/TLR)) RNA. Using this construct well diffracting crystals could be generated and the crystal structure could be solved. The detailed crystallization procedure and structure determination is described in section 3.3.2.2.

In addition to the *ASH1*-E3 element, it would also be interesting to know the structures of other RNA LEs like *ASH1* E1, E2A, E2B, *WSC2* or *EAR1,* which are all localized by She2p in yeast. Since these LEs do not share any sequence or secondary structure similarity, but are all specifically incorporated into SHE complexes, comparing their structural properties could shed light on their general operation mode. Previously, the *ASH1* E2B and *WSC2* LEs were minimized to a length of 21 nt (E2B (21 nt)) and 36 nt (*WSC2*_N (36 nt)), respectively. These shortened LEs were able to bind She2p with K_Ds in the nanomolar range [203] and were therefore chosen to be crystallized. Since initial screens on chemically synthesized E2B (21 nt) and *WSC2*_N (36 nt) did not result in crystals at 21 °C, all subsequent RNA constructs were fused to the previously successfully utilized TL/TLR scaffold.

Since crystallization requires large amounts of highly pure RNA, the following TL/TLR-scaffold constructs were produced by large scale *in vitro* transcription (section 2.10.5). For this purpose, T7 RNA polymerase was bacterially expressed and purified (section 2.9.4). Initial small scale (50 µl) *in vitro* transcription approaches were applied for each construct to screen for the most efficient

magnesium concentration resulting in the highest RNA yield. An exemplary magnesium optimization screen is shown in Figure 3.7. Optimal, construct dependent magnesium concentrations are summarized in Figure 3.8 A. RNAs from large scale *in vitro* transcription reactions (up to 30 ml) were then separated via urea PAGE. The appropriate band was excised with the help of UV shadowing and RNA was extracted by electro elution. After dialysis against 1 M NaCl and twice against RNase-free water RNA was refolded in 10 mM sodium cacodylate pH 6.5, 5 mM $MgCl_2$ and its homogeneity examined by urea and native PAGE (section 2.10.3). Corresponding yields of individual large scale *in vitro* transcriptions of different constructs are also listed in Figure 3.8 A.
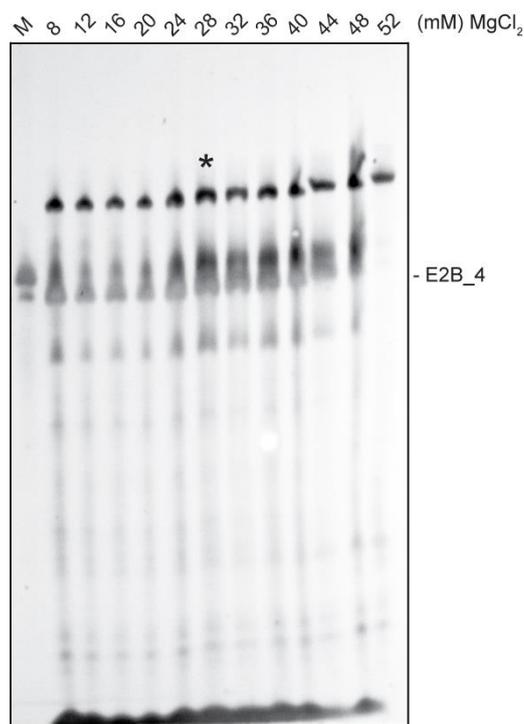
Figure 3.7: Exemplary magnesium optimization screen for large scale *in vitro* transcription reactions. 8 % Midi urea PAGE gel of denatured E2B_4 (56 nt-TL/TLR) samples from 50 µl *in vitro* transcription reactions with the indicated $MgCl_2$ concentration. Asterisk marks the most efficient $MgCl_2$ concentration for construct E2B_4 (56 nt-TL/TLR), which was 28 mM. 200 ng of a 61 nt DNA oligonucleotide served as marker for a rough size estimate.

Construct E2B_2 (40 nt-TL/TLR) consisted of E2B (21 nt), 5' fused to the 17 nt/ GAAA scaffold and 3' lacking two flexible bases (Appendix Figure 5.5 F). Since this RNA did not result in crystal formation, the 12 nt longer E2B_1 (54 nt-TL/TLR) was generated (Appendix Figure 5.5 E). Although this construct yielded plenty and tiny crystals in initial screens at 4 mg/ml, which could be increased in size by refinement screens (Figure 3.8 C), they were still too small for harvesting. Since small changes in base

composition either at the 5' and 3' ends or in the scaffold stem can dramatically change the crystal's appearance (Figure 1 in reference [202]), E2B_3 (56 nt-TL/TLR) and E2B_4 (56 nt-TL/TLR) were designed (Appendix Figure 5.5 G and Appendix Figure 5.6 A, respectively). Indeed, addition of an extra base pair to the scaffold stem changed tiny crystals to long rods (Figure 3.8 D). The extra base pair at the proximal LE stem resulted in uniformly shaped tetragonal crystals (Figure 3.8 E). Most promising crystals were frozen with different cryogenic protectants (20 % and 25 % of ethylene glycol and glycerol) and their diffraction qualities tested at the synchrotron ESRF (Grenoble, France). Albeit RNA crystals containing E2B_3 (56 nt-TL/TLR) did not show any diffraction pattern, crystals consisting of E2B_4 (56 nt-TL/TLR) yielded some spots up to 10 Å resolution. The RNA content of E2B_4 (56 nt-TL/TLR) crystals was confirmed on an agarose gel (Figure 3.8 F).

To assess the E1 LE of *ASH1*, the scaffold construct E1_1 (64 nt-TL/TLR) was designed (Appendix Figure 5.5 A). Like for the crystallized E3-(42 nt-TL/TLR) RNA, part of the distal E1 loop was replaced by the 17 nt/ GAAA crystallization scaffold (Appendix Figure 5.5 A). By using the pre-crystallization test (Hampton Research), a concentration of 4.3 mg/ml was determined to be well suited for crystallization. Unfortunately no crystals grew in initial trials. Therefore E1_2 (68 nt-TL/TLR) was created (Appendix Figure 5.5 B). In E1_2 (68 nt-TL/TLR) two heterologous base pairs were added proximal to the scaffold stem to probably alter the crystallization behavior of the construct. Using this RNA, at least tiny micro crystals grew in few conditions but were not followed-up.

The E2A (65 nt-TL/TLR) RNA, which carried the scaffold 5' (Appendix Figure 5.5 C), yielded spherulites and microcrystals in several initial conditions after one week at 7 mg/ml and 21 °C. Since it is known that additives sometimes improve the quality and size of macromolecular crystals by changing sample-sample or sample-solvent interactions, most promising conditions of misshaped and tiny crystals (50 mM Na cacodylate pH 6.5, 10 mM MgOAc, 1.3 M $LiSO_4$ and 50 mM Na cacodylate pH 6.5, 200 mM KCl, 10 % (w/v) PEG 8000, 100 mM MgOAc, respectively) were subjected to additive and heavy atom screening (section 2.12.5). Although the addition of 25 mM ammonium oxalate improved the crystal's appearance (Figure 3.8 B), they were still too tiny to be harvested.

When testing the *EAR1* (71 nt-TL/TLR) and *WSC2*_N (69 nt-TL/TLR) LEs, which were 5' fused to the 17nt/GAAA scaffold (Appendix Figure 5.7 A and C, respectively), no crystals could be obtained.

A

Name		Scaffold	Secondary structure predictions	$c_{(MgCl2\ opt.)}$ for *in vitro* transcription	Yield per 10 ml transcription reaction	Crystals
ASH1						
E1_1	(64 nt-TL/TLR)	17 nt/ GAAA	Figure 5.5 A	32 mM	1.15 mg	-
E1_2	(68 nt-TL/TLR)	21 nt/ GAAA	Figure 5.5 B	16 mM	2.79 mg	-
E2A	(65 nt-TL/TLR)	17 nt/ GAAA	Figure 5.5 C	44 mM	1.66 mg	+, see B
E2B	(21 nt)	-	Figure 5.5 D	-	-	-
E2B_1	(54 nt-TL/TLR)	17 nt/ GAAA	Figure 5.5 E	28 mM	0.5 mg	+, see C
E2B_2	(40 nt-TL/TLR)	17 nt/ GAAA	Figure 5.5 F	28 mM	0.4 mg	-
E2B_3	(56 nt-TL/TLR)	19 nt/ GAAA	Figure 5.5 G	28 mM	0.3 mg	+, see D
E2B_4	(56 nt-TL/TLR)	17 nt/ GAAA	Figure 5.6 A	28 mM	0.4 mg	+, see E, F
E3	(25 nt-loop)	-	Figure 5.6 B	-	-	-
E3	(28 nt-loop)	-	Figure 5.6 C	-	-	-
E3	(42 nt-TL/TLR)	17 nt/ GAAA	Figure 5.6 E	-	-	+, see section 3.3.2.2
EAR1						
EAR1	(71 nt-TL/TLR)	17 nt/ GAAA	Figure 5.7 A	36 mM	0.3 mg	-
WSC2						
*WSC2*_N (36 nt)		-	Figure 5.7 B	-	-	-
*WSC2*_N (69 nt-TL/TLR)		17 nt/ GAAA	Figure 5.7 C	40 mM	0.8 mg	-

B **C** **D** **E** **F**
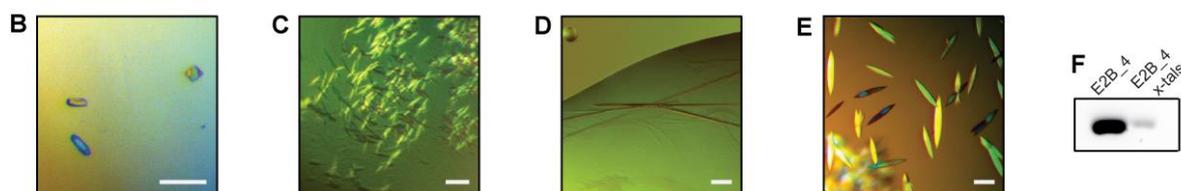
Figure 3.8: Overview of different RNA constructs used for LE-crystallization trials. A) Table summarizing all crystallized RNA constructs with and without donor/ acceptor scaffold. Secondary structure predictions can be found in Appendix **Figure 5.5**, **Figure 5.6** and **Figure 5.7**. For highest *in vitro* transcription yields, optimal magnesium concentrations were determined. Respective 10 ml transcription yields are listed. Crystals of **B)-D)** Crystals were growing at room temperature using the hanging drop vapor diffusion technique [191]. Depicted scale bars represent 100 μm. **B)** E2A (65 nt-TL/TLR) grown in 50 mM Na cacodylate pH 6.5, 200 mM KCl, 10 % (w/v) PEG 8000, 100 mM MgOAc and 25 mM ammonium oxalate. c_{RNA} = 7 mg/ml; **C)** E2B_1 (54 nt-TL/TLR) grown in 50 mM Tris, pH 8.5, 2.7 M hexanediol, 5 mM MgSO₄. c_{RNA} = 2.5 mg/ml; **D)** E2B_3 (56 nt-TL/TLR) grown in 40 mM Na cacodylate pH 5.5, 80 mM KCl, 12 mM NaCl, 40 % (v/v) MPD, 20 mM hexamine cobalt(III)chloride. c_{RNA} = 2.5 mg/ml; **E)** E2B_4 (56 nt-TL/TLR) grown in 50 mM Na cacodylate pH 6.0, 100 mM NaCl, 10 % (w/v) PEG 4000. c_{RNA} = 2.5 mg/ml. **F)** 0.5 % agarose gel with the denatured E2B_4 (56 nt-TL/TLR) RNA sample before crystallization (300 ng) and after extraction from crystals of condition D), showing that it's indeed crystallized RNA and not salt.

3.3.2.2 Crystallization and structure determination of the E3-localization element of the ASH1 mRNA

The chemically synthesized E3 (42 nt-TL/TLR) RNA (Figure 3.9 A, Appendix Figure 5.6 E, sequence see Table 2.5) was deprotected and refolded as described in section 2.12.3. Initial crystallization trials were set at a concentration of 6.98 mg/ml at 21 °C and yielded crystals in several conditions.

After varying pH, precipitant and RNA concentration, best crystals grew in 50 mM MES pH 6.0, 5 % (w/v) PEG 4000, 5 mM MgSO$_4$ by using the hanging-drop vapor diffusion method [191] with 500 µl reservoir volume in EasyXtal 15-Well plates (Qiagen). Mixing 1 µl well solution and 1 µl RNA (2.94 mg/ml) at 21 °C yielded tetragonal crystals of approximately 125 x 44 x 44 µm size after three days (Figure 3.9 B). Prior to data collection the crystals were cryogenically protected for 30 s in well solution, supplemented with 25 % (w/v) PEG400 and flash-frozen in liquid nitrogen.

Native data set was recorded at 100 K at beam line ID 23-2 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). Data were indexed and integrated using XDS and scaled via XSCALE [204]. Structure factor amplitudes were obtained via the Truncate program (CCP4 package [195]). Phases could be gained by molecular replacement with Phaser [205] using the slightly modified PDB entry 4FNJ (structure of the tetraloop-receptor) as search model. Crystals, diffracting up to 2.65 Å resolution (Figure 3.9 C), had the space group $P4_12_12$ and one RNA molecule per asymmetric unit (see also Figure 3.9 D). The structure was completed by manual building with the program COOT [206], followed by several rounds of refinement with the program PHENIX [207]. All crystallographic software was used as part of the SBGrid software distribution [208]. The structure is deposited at the protein data bank (PDB-ID: 5M0H).

The final structure (Figure 3.9 E, F) showed an elongated RNA stem, where three bases (C1779, A1782 and A1783) were flipped out and pointed to the solvent. These bases showed very little electron density and are therefore probably flexible. The integrity of the RNA in the crystal was confirmed by agarose gel (Figure 3.9 G).

A

E3
(28 nt-loop)

E3
(42 nt-TL/TLR)

B

C

D

Data collection/ processing	
Beam line	ESRF ID 23-2
Wavelength (Å)	0.8726
Detector distance (mm)	318.431
Number of images	47
Oscillation range (°)	2
Space group	$P4_12_12$
Cell dimensions	
a, b, c (Å)	64.32, 64.32, 75.19
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	50-2.65 (2.72-2.65)*
R_{sym} or R_{merge}	7.3 (107.2)
I/σI	19.93 (2.27)
Completeness (%)	99.6 (100.0)
Redundancy	7.24 (7.5)

Refinement	
Resolution (Å)	45.48-2.65
No. of reflections	4,920
R_{work} / R_{free} (%)	21.1 / 24.0
No. of atoms	
RNA	894
Ion	5
Water	10
B-factors	
RNA	91.9
Ions	49.0
Water	67.7
R.m.s. deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.21

E

F

C1779
A1782
A1783

A1783
A1782
C1779
5'
3'

G

E3 (42 nt-TL/TLR)
dissolved crystals

Figure 3.9: Crystallization and structure determination the minimal E3-localization element. A) Secondary structure predictions of the E3 (28 nt-loop) RNA (left) and the E3 (42 nt-TL/TLR) RNA (right). The green part consists of E3 nucleotides 1774-1811, in which positions 1786-1802 had been replaced by a 17 nt acceptor GAAA-receptor-containing hairpin (grey). **B)** Representative RNA crystal grown in 50 mM MES, pH 6.0, 5 % (w/v) PEG 4000, 5 mM MgSO$_4$. White scale bar represents 100 μm. **C)** Diffraction image of the big crystal from B). **D)** Data collection/ processing and refinement statistics (molecular replacement) *Values in parentheses are for highest-resolution shell. **E)** Crystal structure of the E3 element (green), and the heterologous hairpin (gray) at 2.65 Å resolution. In the stem bases C1779 and 1782 to 1783 are pointing towards the solvent. PDB-ID: 5M0H. **F)** Bottom view of E). **G)** Agarose gel with the denatured RNA sample before crystallization and after extraction from crystals indicates no degradation.

3.3.3 NMR analysis of the *ASH1*-E3 element indicates conformational flexibility

Since the crystallized RNA consisted only in part of the E3 element and contained a rigid heterologous stem (Figure 3.9 A, E), NMR experiments should clarify the question whether the base pairing of E3 (42 nt-TL/TLR) is comparable with the base pairing of E3 (28 nt-loop) in solution. For this purpose the optimal magnesium concentrations of E3 (42 nt-TL/TLR) and E3 (28 nt-loop) RNAs for *in vitro* transcription were determined (20 mM and 24 mM, respectively), and the RNAs produced and purified in large scale (section 2.10.5). NMR experiments were performed and analyzed by Dr. Andreas Schlundt (Chair of Biomolecular NMR Spectroscopy, Department of Chemistry, TUM, Munich, Germany) (section 2.12.2).

After refolding the RNAs, 1D-imino proton spectra were recorded. Additional 2D-imino NOESY experiments were used to sequentially assign the observable imino protons (Appendix Figure 5.9 A, B). These constitute an unambiguous readout for RNA base pairing, as imino protons are only observable in stable H-bonds. Comparable regions of E3-(42 nt-TL/TLR) and E3-(28 nt-loop) RNAs showed very similar base pairing in solution. In both cases, the lower and upper stems enclosing the central double bulged region (Figure 3.9 A) form stable base pairs, whereas slight differences were visible in this central bulged region (Appendix Figure 5.9 C). In E3 (42 nt-TL/TLR) imino-proton signals for U1780 and U1784 were not detectable and the signal corresponding to G1781 was significantly line-broadened, likely indicating conformational dynamics. In contrast the corresponding imino signals for U1780 and G1781 in E3-(28 nt-loop) RNA were clearly detectable. Here the GAAA tetra-loop, which is located in close proximity to U1780 and G1781 probably lead to stabilization in this region. Nevertheless, together with the imino resonances of U1807 all three signals showed extreme line-broadening, indicating flexibility and dynamics in the central double bulged area.

Overall these data indicate that the base pairing observed in the crystal structure (Figure 3.9 E) is very similar in solution. Although the E3-(42 nt-TL/TLR) RNA shows some flexibility and melted base-pairs in the central double bulged region, the narrow imino signal for U1807 and the detection of the G1781 imino indicate that a significant fraction of the E3-(42 nt-TL/TLR) RNA appears in an elongated conformation as seen in the apo-crystal structure. Notably, the complete set of imino correlations and thus base-pairing was only observed in the E3 (28 nt-loop) RNA.

In summary, NMR data show that an elongated state likely represents a major conformation of the E3 RNA in solution. Considering the dynamics in the central double bulged region the free RNA is probably able to adopt different conformations.

3.3.4 Crystallizing She2p in co-complex with RNA

The molecular understanding how the RNA-binding protein She2p accomplishes RNA recognition remains elusive to date. In former times, there have already been several trials on crystallization of She2p together with RNA constructs deriving from different *ASH1*-localization elements. Unfortunately the obtained crystals did not diffract better than 9 Å [203]. My attempt to reproduce those crystals consisting of She2p (6-246) together with E3 (41 nt) or with E3 (28 nt-loop) failed, although broad condition screening was applied. Therefore new initial screening was set up.

3.3.4.1 Screening for crystallizable She2p-RNA combinations

Since a quadruple cysteine mutant of She2p yielded well-diffracting crystals [150], the cysteine-mutated version of She2p (She2p (6-246, C-S)) in complex with the E3 (28 nt-loop) RNA was utilized. It turned out that the interaction of She2p and E3 RNA was too weak to purify the co-complex by size-exclusion chromatography, therefore protein and RNA was mixed at a molar ratio of 2:1.2 in a buffer containing 20 mM Hepes/NaOH pH 7.8, 200 mM NaCl and 2 mM $MgCl_2$ and directly crystallized. Crystals grew slowly in several conditions both at 4 °C and 20 °C after approximately one month. When either She2p or RNA was omitted from the crystallization set-up no crystals were growing, indicating that indeed the She2p:E3 co-complex was crystallized. Further hanging-drop refinement of promising crystals included varying the complex concentration, drop-size, pH and PEG concentration. Best crystals that grew from precipitate in 100 mM BIS Tris propane pH 6.5, 200 mM sodium acetate, 20 % (w/v) PEG 3350, were frozen in 25 % ethylene glycol and diffracted up to 3.7 Å resolution. Due to suboptimal diffraction the data set had to be limited to 4.5 Å resolution after processing. She2p:RNA crystals were monoclinic and had the same *C*2 space group as She2p [150], but altered unit cell parameters and more molecules (8 x She2p) per asymmetric unit. By using the published She2p crystal structure (PDB-ID: 1XLY) [150] for molecular replacement with the program MolRep [209] and after several rounds of refinement with the program RefMac5 [210], [211], additional unassigned electron densities were visible that clearly suggested the presence of RNA. Unfortunately RNA could not be modeled into the densities, raising the necessity to improve crystals further.

In this respect, crystals were reproduced and tested with different cryogenic protectants but without any improvement in diffraction quality. Although in an heavy atom screen (section 2.12.5) no crystals were growing, an additive screen (section 2.12.5) yielded two promising conditions additionally containing either 3 % (v/v) MPD or 4 % (v/v) 1-propanol. The latter additive yielded small, needle-

shaped crystals, which could be reproduced but showed just few diffraction spots up to 3.3 Å resoltuion. Since in case of She2p alone [150] heavy atom soaking resulted in improved diffraction quality, She2p:RNA crystals were soaked with K_2PtI_6 (section 2.12.11). Tested crystals did not show any diffraction better than 8 Å.

The diffraction quality of She2p-only [150] increased dramatically when crystals were generated with C-terminally truncated She2p mutants. For this reason not only She2p (6-246, C-S):E3 (28 nt-loop), but also co-complexes together with She2p (6-243, C-S) or She2p (6-240, C-S) were set up in the best diffracting condition, both at 4 °C and 20 °C and in two different drop sizes. Since in these trials no crystals were growing She2p (6-246, C-S) was again used further on.

Because crystals grew in several conditions with sodium and varying counter ions, different refinement screens including for example 200 mM sodium sulfate or formate were set up. Crystal measurement at ESRF revealed an improved diffraction of She2p:RNA up to 3.3 Å resolution in 100 mM BIS Tris propane pH 6.5, 200 mM sodium fluoride and 20 % (w/v) PEG 3350. In this case crystals grew faster and were already visible after 10 days. The resulting symmetry (*C*2) was again the same but with slightly changed unit cell parameters and only four She2p molecules per asymmetric unit. Additional densities for two RNA molecules per She2p tetramer were visible in close proximity to amino acids forming the basic helical hairpin [150]. For She2p some additional amino acids could be assigned, which were not visible in the initial structure. Nevertheless neither trying to build single bases of E3 (28 nt-loop) into the density, nor using molecular replacement of parts of the solved RNA structure (see section 3.3.2.2) were successful to solve the co-complex structure. By just putting phosphor atoms to the strongest densities, the quality of the electron density map improved to an extent where structure solving was possible. Nevertheless, after several rounds of refinement, the quality of the RNA density still left several conformational uncertainties, raising the necessity to improve the diffraction quality of the crystals even further.

Since reducing the solvent content of the crystals was reported to improve the diffraction quality dramatically in some cases [212], She2p:RNA crystals were subjected to dehydration (section 2.12.8). Beneath using either oil or 30 % PEG 3350 as cryogenic protectant, also varying exposure times with each of these cryogenic protectants were tested. The best diffracting crystal showed diffraction up to 3.9 Å resolution. Since these crystals were plate-like, very thin and mostly small, an additive and heavy atom screen (section 2.12.5) on She2p:E3 was applied to improve the crystal growth. Few and slightly more voluminous crystals grew by adding 10 mM $NiCl_2$, but were not followed up further.

In course of crystallization trials on the ternary complex consisting of She2p, She3p and RNA, which were performed in parallel I soaked She2p:RNA crystals with a 24 aa She3p peptide (residues 382-405). This approach yielded finally a data set with improved diffraction up to 2.2 Å resolution, which had to be limited to 2.4 Å resolution after processing. Here the assignment of free electron density was finally unambiguously possible. The detailed crystallization procedure and structure determination is described in the next section 3.3.4.2.

An additional idea was to crystallize She2p together with other RNA LEs like *ASH1*-E2B or parts of *WSC2*-N. Therefore initial crystallization trials of She2p (6-246, C-S) and either E2B (21 nt) or *WSC2*-N (36 nt) were set up at room temperature and 4 °C. Most promising conditions were subjected to refinement by screening for different precipitant and salt concentrations as well as varying the pH (section 2.12.4). By using an automated liquid-handling station (Tecan Freedom EVO®) a broad grid of conditions was screened. The most promising She2p:E2B crystals grew in 100 mM MES pH 6.5, 1.6 M $MgSO_4$ exclusively at 4 °C and diffracted up to 6.0 Å. A full data set was collected, but data processing was not successful. For She2p:*WSC2*-N best crystals grew in 100 mM Tris pH 8, 200 mM $CaCl_2$ and 20 % (w/v) PEG 6000. Trials to reproduce those crystals failed and were not pursued further.

3.3.4.2 Crystallization and structure determination of the She2p-ASH1 E3 co-complex

Crystallization conditions of She2p together with RNA were initially screened at 4 °C using a Mosquito nanodrop dispenser (TTP Labtec) in 96-well sitting-drop plates and commercial screens. 4 mg/ml She2p (6-246, C-S) were mixed with 0.77 mg/ml E3 (28 nt-loop) RNA (sequence see Table 2.5) in a final buffer containing 20 mM Hepes, pH 7.8, 200 mM NaCl and 2 mM $MgCl_2$. The complex was assembled for 10 min at RT, spun down 16,100 x g, 4 °C, 60 min. Thin and plate-shaped crystals could be reproduced via hanging-drop vapor diffusion refinement, by equally mixing 1 μl co-complex with 1 μl 100 mM Bis Tris propane, pH 6.5, 200 mM NaF, 20 % (w/v) PEG 3350, at 4 °C. The reservoir volume in EasyXtal 15-Well plates (Qiagen) was 500 μl. Native crystals, slowly growing over a period of three weeks, were soaked with 1 mM She3p peptide. Therefore, commercially synthesized, HPLC-purified She3p (382-405) (JPT Peptide Technologies) was solubilized in water and dialyzed against crystallization condition over night, 4 °C, using a dialysis button (Hampton Research) and a 1,000 Da cut-off ZelluTrans membrane (Roth). Crystals were cryogenically protected by 25 % ethylene glycol being added to the crystallization condition and flash-frozen in liquid nitrogen.

Data collection at 100 K was performed at beam line ID 23-1 at the synchrotron ESRF (Grenoble, France). Data were indexed, integrated and scaled using the XDS program package [204]. Structure factor amplitudes were obtained via the Truncate program [195]. Initial electron-density maps were calculated by molecular replacement with the apo-structure of She2p (PDB-ID: 1XLY) as search model using MolRep (CCP4 program suite). RNA and She3p were manually built using the program COOT [206]. The overall structure at 2.41 Å resolution was refined using the program RefMac 5.8 [210, 211], including non-crystallographic symmetry (NCS) averaging, the maximum-likelihood target function with translation-liberation-screw (TLS) parameters [213] and the PDB_REDO pipeline (version 6.0, [214]). The final model is characterized by R_{work} and R_{free} factors of 18.9 % and 23.5 %, respectively. Data collection, processing and refinement parameters are listed in Appendix Table 5.2. The final structure is deposited at the protein data bank (PDB-ID: 5M0I).

3.3.4.3 Crystal structure of the nuclear complex consisting of She2p and the ASH1-E3 element

The structure was solved and refined as described in section 3.3.4.2. One unit cell of the crystal lattice consists of four She2p (6-246, C-S) and two E3 (28 nt-loop) RNA molecules. Parts of the She3p peptide (382-405) that were needed to improve the diffraction quality were present three times. Nevertheless its binding to She2p was determined to be not physiologic (for details see section 3.3.4.5) for which reason it is not depicted in the co-complex structure (Figure 3.10).

Two She2p homodimers sit on top of each other in a head-to-head conformation and form an elongated, two-fold point symmetric tetramer (Figure 3.10). In the co-complex the tetramer shows no major conformational changes when comparing the She2p-C_α atoms to its equivalents in the unbound state (root-mean-square deviation (RMSD) of PDB-ID 1XLY = 0.74 Å and of PDB-ID 4WNL = 0.97 Å). Two RNA molecules (E3 and E3') that adopt a kinked conformation bind to the positively charged surface areas between the protruding helices on opposing sides of the tetramer. Apart from the very 3' base that does not make any protein contacts (Figure 3.11 A) both RNA molecules are conformational identical (RMSD of all atoms = 0.5 Å). Interestingly the RNA has to undergo large conformational changes to fit between the protruding helices of She2p. Details on this RNA rearrangement are described in section 3.3.4.4. A striking feature in the co-structure is the point-symmetry of the She2p tetramer, whereas the RNA itself is asymmetric.
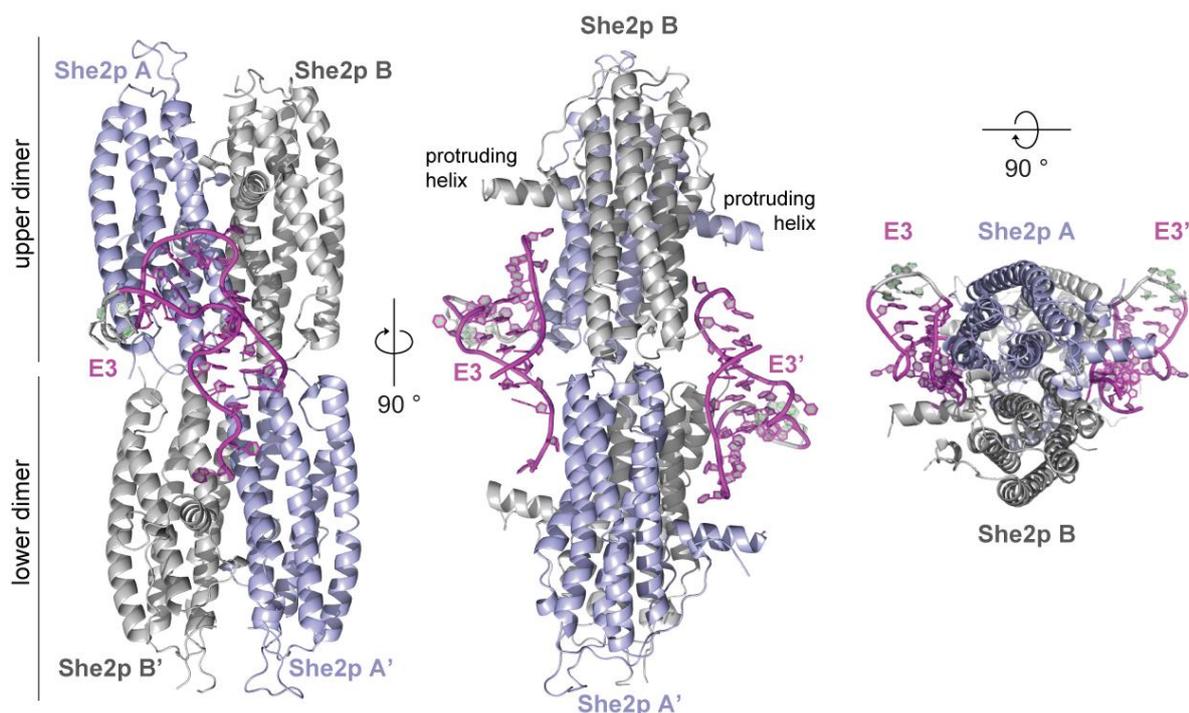
Figure 3.10: Structure of She2p in complex with the minimal *ASH1* E3-localization element. Crystal structure of She2p (6-246, C-S) and *ASH1*-E3 (28 nt-loop) RNA at 2.41 Å as cartoon representation. Front view (left): Two She2p homodimers form an elongated tetramer with a non-crystallographic two-fold point symmetry. Monomers are depicted in light blue (A and A') and gray (B and B'). Side view (middle): The symmetric tetramer binds two RNA molecules (E3 and E3', pink) in a kinked conformation on opposing sides between the protruding helices in an asymmetric way. The heterologous GAAA tetra-loop in E3 and E3' is shown in light gray. Top view (right) of the complex illustrates the asymmetric binding of E3 and E3' to the symmetric tetramer.

The RNA is bound by a variety of different She2p amino acids (Figure 3.11 A). Although one of the two RNA chains has fewer contacts to She2p, the majority of interactions is identical. Interestingly one E3 molecule contacts three different She2p monomers on each side of the tetramer. The conserved cytosines C1779 and C1813 show the strongest coordination. C1779, located in the kink of E3, is contacted by N36, R63 via base-specific interactions and by R52 via backbone interactions (Figure 3.11 B) in the upper She2p dimer. The very same set of interactions can be observed in the lower dimer for C1813 residing at the 3' end of E3 (Figure 3.11 C). In summary two sets of N36, R63 and R52 residues on each side of the She2p tetramer specifically recognize these two cytosines in each RNA and act as rulers for their correct distance (37-38 Å between their ribose C1' positions). These observations explain why single amino-acid mutations of either N36, R63 or R52 lead to a complete loss of *ASH1*-mRNA binding and defective localization *in vivo* and abrogated RNA binding of She2p *in vitro* [150], [151].

Upon co-complex formation U1780 becomes unpaired, flips away and base stacks with F176 in the protruding helix of She2p. The adjacent E172 helps to position the base (Figure 3.11 D). The second protruding helix on this side of the tetramer fails to contact the RNA since a guanine (G1814) instead of an uracil is located in this respective position (Figure 3.11 E). The importance of this protruding helix is underlined by the fact that a deletion mutant of this helix leads to abrogated *ASH1*-mRNA localization *in vivo* [139].
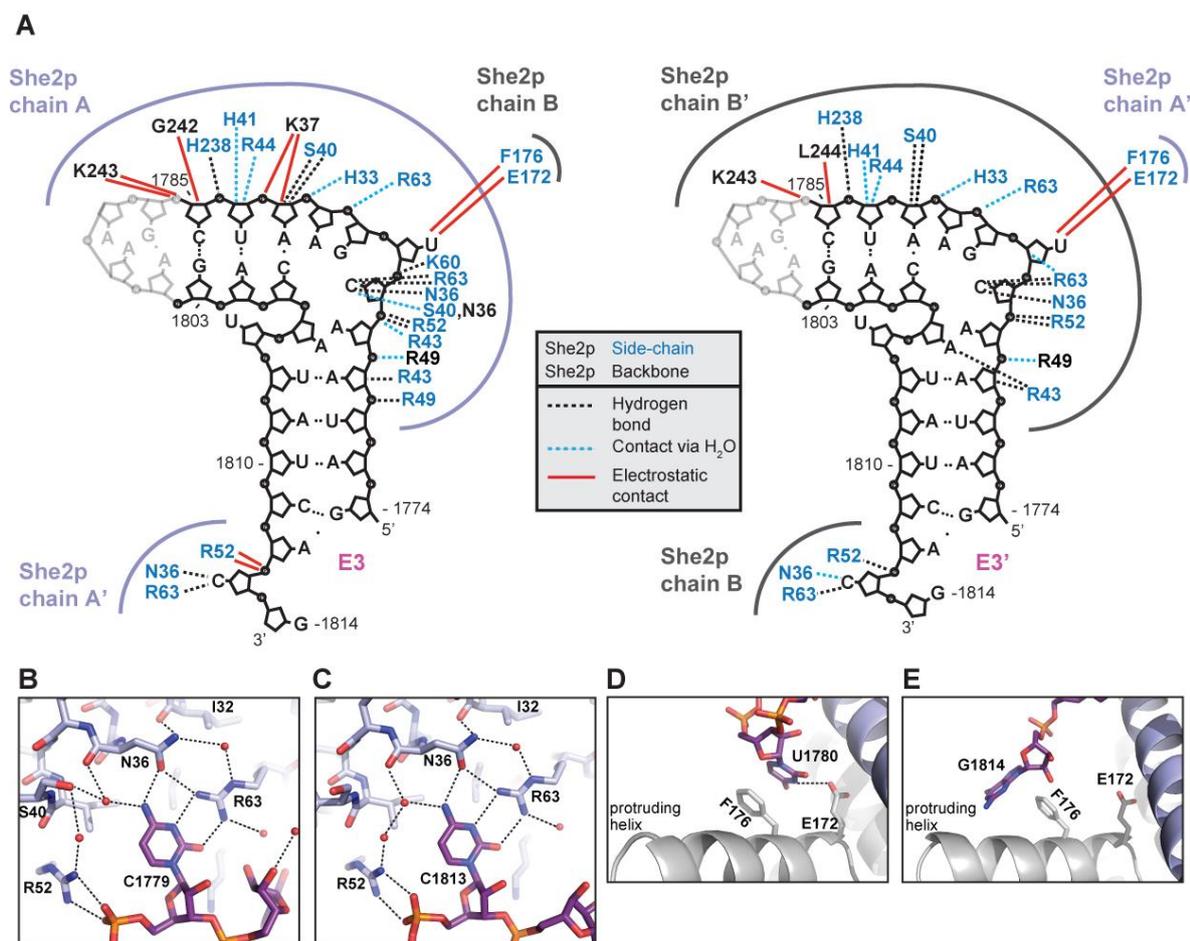
Figure 3.11: Details of *ASH1* E3-RNA recognition by She2p. A) Schematic representation of She2p-*ASH1* E3-RNA contacts. Each RNA in the co-complex E3 (left) and E3' (right) contacts three different She2p chains of the She2p tetramer. Color-coding of the different chains is equal to **Figure 3.10**. Protein-RNA contacts derive from Nucplot [196] analysis. **B)** and **C)** show close-ups of C1779 and C1813 from E3 of **Figure 3.10**, respectively. The conserved cytosine C1779 is highly coordinated by N36 and R63 of chain A from the upper She2p dimer. The backbone of I32 helps to arrange N36 properly. S40 contributes to the binding via hydrogen bonding. R52 fixes the phosphate position of C1779. The second conserved cytosine C1813 in C) shows the same contacts, but to chain A' of the lower She2p dimer. Sole difference is the missing interaction of S40 due to an alternative rotamer. **D)** Upon co-complex formation U1780 flips out and base stacks with F176 of the protruding helix from chain B of the upper She2p dimer. An electrostatic contact of E172 contributes to arrange U1780. **E)** Due to sequence deviation at the 3' end of the RNA (G instead of U) base 1814 is not able to interact with F176 from chain B' of the lower She2p dimer.

She2p contacts the RNA also with its very C-terminus. Here amino acids H238, K243 and either G242 or L244 in E3 or E3', respectively, contribute to RNA binding at one side of the distal stem (Figure 3.11 A). In line with this observation, deletion of the She2p C-terminus showed reduced RNA-binding capabilities in filter-binding assays [139].

3.3.4.4 Dynamics in the E3 LE bulge-region indicates a conformational sampling mode for the ASH1-E3 RNA-recognition

In line with NMR experiments that revealed a flexible part in the central bulged region of the unbound RNA (section 3.3.3), bases in this area have to undergo tremendous conformational changes to kink the RNA and to fit She2p (Figure 3.12 and Appendix Figure 5.10).
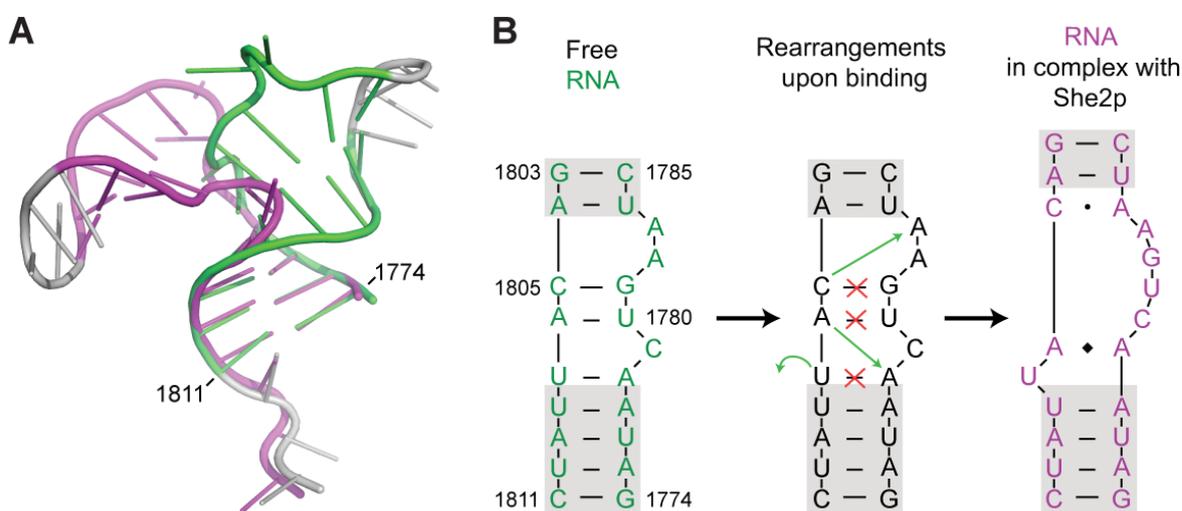
Figure 3.12: Conformational changes of the E3 LE upon co-complex formation with She2p. A) Overlay of an RNA model deriving from the crystal structure of the E3 element (green) and of the E3 (28 nt-loop) RNA from the co-complex with She2p (pink). In the green RNA the heterologous TL/TLR scaffold has been replaced *in silico* by a GAAA tetra-loop (gray) and the last three bases at the 3' end (gray) were also manually added. B) Schematic visualization of secondary structure base rearrangements in the central bulged region of the E3 LE upon She2p binding. While the upper and lower stem (shimmed in gray) stays unaffected, several bases in the bulged region lose their base pairing (indicated by red crosses) and either stay unpaired or interact with another base instead (indicated by green arrows). Diamond indicates that A1806 interacts via its ribose with the base of A1778. For more details on internal RNA interactions see **Appendix Figure 5.10**.

To allow for these rearrangements part of the double stranded region either changes its base pairing or becomes unpaired. For instance A1783, which pointed to the solvent in the free RNA (Figure 3.9 E in section 3.3.2.2) establishes a new contact to C1805 on the opposing side (Figure 3.12 B). As a result the G1781-C1805 hydrogen bonds are broken up and G1781 undergoes base stacking with

A1782. Additionally the base pairing of U1780 and A1806 is abrogated upon co-complex formation and U1780 flips away, making new contacts to E172, F176 of the protruding helix in She2p (Figure 3.11 A, D). Furthermore the connection between A1778 and U1807 breaks and instead a stabilizing interaction between the base of A1778 and the sugar part of A1806 is established. Base stacking of both adenines strengthens this interaction. As a consequence U1807 flips away and points towards the solvent. The resulting kinked RNA conformation is further stabilized via several contacts along the positively charged surface region of She2p (section 3.3.4.3 Figure 3.10 and Figure 3.11 A). In line with NMR experiments these results indicate a conformational sampling mode for the *ASH1*-E3 RNA-recognition.

3.3.4.5 Binding of She3p (382-405) to She2p is not physiologic

As already mentioned, soaking She3p (382-405) to the She2p-RNA crystal was necessary to improve its diffraction quality and to unambiguously solve the co-complex structure (section 3.3.4.2 and 3.3.4.3). When inspecting the She2p-RNA structure three additional electron-density stretches were observed that could neither be assigned to She2p nor to RNA (Figure 3.13 A).

However, the position of this density overlapped with a previously published interaction site of She3p (364-368) [153]. Therefore both sequence regions of She3p comprising either (382-405) or (364-368) were aligned. Interestingly, sequence similarities could be observed between residues 400-403 and 364-368 (Figure 3.13 B).

By superposing the She2p-She3p structure [153] on the She2p-RNA co-complex amino acids 398-405 could be modeled into the extra electron density (Figure 3.13 C).

To assess the physiologic relevance of this peptide-protein interaction ITC experiments were performed (Figure 3.13 D). She3p (382-405) showed no binding to She2p wild type even at a final concentration of 157 µM, while the published She3p (342-374)Δ(352-354) peptide [153] did. When also considering that the published MBP-tagged, She3p (342-374)Δ(352-354) peptide had a reasonable affinity of 2.45 µM for She2p I can conclude that the electron density of She3p (382-405) in the co-structure of She2p and RNA is non-physiologic.
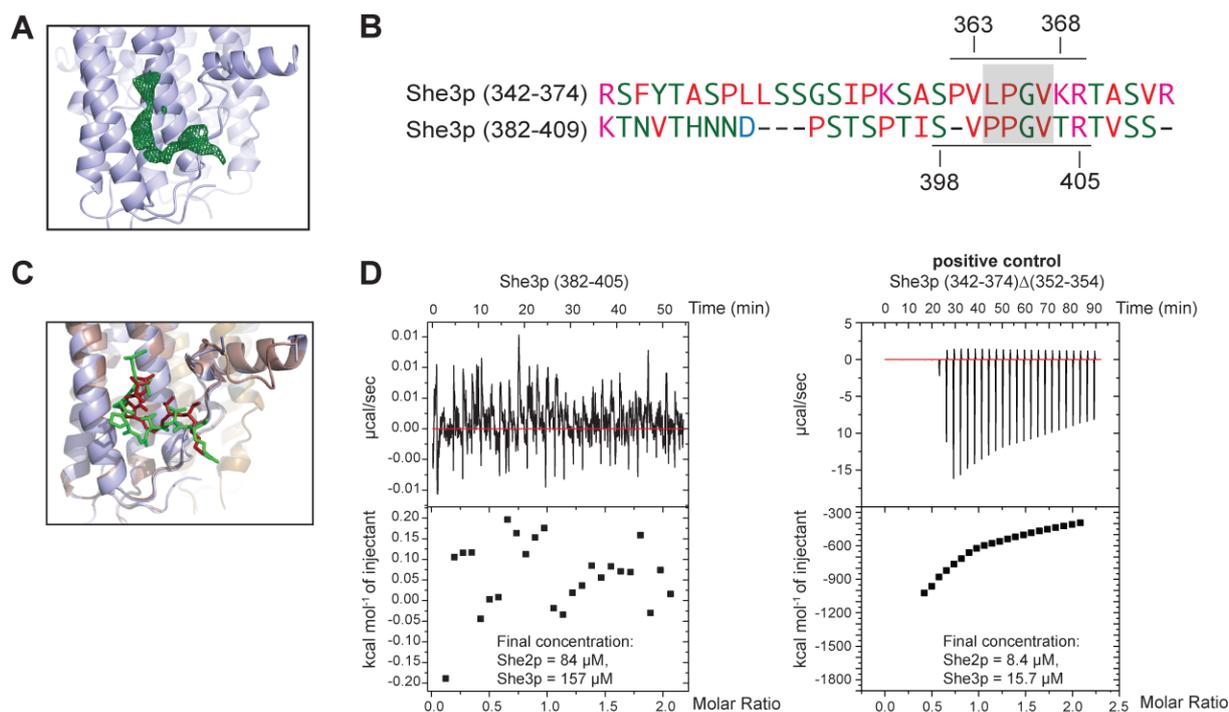
Figure 3.13: Binding of She3p (382-405) to She2p is not physiologic. A) Section of the She2p-RNA co-complex of **Figure 3.10**, showing exemplarily one of three additional electron densities (green) that could neither be assigned to She2p nor to RNA. **B)** Alignment of two different She3p regions shows sequence similarities. She3p (382-405) was used for She2p-RNA co-complex crystallization. She3p (342-374) comprises the previously published 'LPGV' motif [153] known to bind She2p (similar region between both is shimmed in gray). **C)** The previously published She2p (copper) and its bound She3p peptide (363-368, red) is superposed with She2p (light-blue) from the co-complex structure. Residues (398-405, green) of She3p (382-405) could be modeled into the electron density shown in A). Protein-protein contacts are similar, but not identical. **D)** ITC experiments show no binding of She3p (382-405) to She2p wild type even at a final She3p concentration of 157 µM (left). In contrast, the positive control She3p (342-374)Δ(352-354) [153] showed binding to She2p.

3.3.4.6 SAXS experiments confirm E3 kinking upon She2p binding

In addition to NMR studies (section 3.3.3), SAXS experiments of the unbound *ASH1*-E3 LE were performed to gain insights into the conformational composition of the RNA in solution. For this purpose, E3-(28 nt-loop) RNA was produced in large scale by *in vitro* transcription (section 2.10.5) and SAXS measurements (section 2.12.1) were performed together with Dr. Ralf Stehle (Chair of Biomolecular NMR Spectroscopy at the Department of Chemistry, TUM, Munich, Germany), who also evaluated the results.

The initial experimental set-up with different RNA concentrations (7, 11 and 20 mg/ml) showed non-linear Guinier plots at low q^2 values that pointed to sample aggregation. To improve data quality the

experiment was repeated with less concentrated RNA and scattering curves were recorded at 1.5 mg/ml. To judge the conformational composition in solution RNA models were used to calculate scattering curves, which were fitted with the experimental curves. Since initial fits for single RNA species, resembling either the elongated (bound) or the kinked (unbound) form, yielded very high χ^2 values, the theoretical scattering curves were calculated from a mixture of (up to three) RNA models. Different combinations of the elongated, kinked or a single-stranded variant of the E3 (28 nt-loop RNA) were used to fit the scattering curve (Appendix Figure 5.11).

The possibility that only an elongated and a kinked RNA conformation is present in solution was shown to be unlikely because of a high χ^2 value of 7.6 (ratio elongated : kinked = 0.72 : 0.27, Appendix Figure 5.11 A). The combination of an elongated and a melted, single-stranded RNA species existing in solution improved the χ^2 value to 1.7 (ratio elongated : single-stranded = 0.65 : 0.34, Appendix Figure 5.11 B). When testing a kinked and a single-stranded RNA model the fit got worse and χ^2 increased to 3.9 (ratio kinked : single-stranded = 0.74 : 0.26, Appendix Figure 5.11 C). The best fit could be achieved by assuming that all three species are present in solution. Here χ^2 improved to 1.6 (ratio kinked : single-stranded : elongated = 0.12 : 0.33 : 0.55, Appendix Figure 5.11 D).

Taken these data together, SAXS experiments substantiate the observation that the E3 (28 nt-loop) RNA has a predominantly elongated conformation in solution. The fact that a single-stranded species is additionally required to fit the curve suggests that an ensemble of RNA conformations exists in solution. However, the kinked species does not seem to reflect a major population and its conformation must therefore be sampled by She2p binding.

3.3.5 Crystallizing She2p-She3p and the RNA

The specificity-mediating features in the nuclear co-complex consisting of She2p and the *ASH1*-E3 LE could be clarified by solving its crystal structure (section 3.3.4.3). Nevertheless the fact that cargo specificity is increased in the ternary complex consisting of She2p, She3p and the *ASH1* mRNA [139] has not been structurally substantiated yet. Therefore the following section will deal with the question: What are the structural determinants that mediate the increase of cargo specificity upon joining of She3p in the cytoplasm?

3.3.5.1 Screening for crystallizable combinations of She2p-She3p and RNA

As evident from a previous study [139], She3p (334-425)His$_6$ was the shortest She3p fragment known to form a stable ternary complex together with She2p and *ASH1*-E3 (51 nt). UV-crosslinking experiments revealed the C-terminal She3p regions 334-340 [139] and 382-405 (unpublished data from Dr. Roland Heym) to be involved in E3 (51 nt)-RNA binding. Based on this information, the sequence alignment of She3p from different yeast strains (Figure S13 in reference [139]) and results from limited proteolysis (section 3.3.5.2) several C-terminal She3p constructs were designed and different *SHE3* segments cloned in pGEX-6P-1 or pETM-43 (Table 2.3). Purification protocols for She3p C-terminal fragments were established and are described in section 2.9.8. Protein expression and purification characteristics are summarized in Table 3.1.

She3p	Affinity tag		Soluble expression	Stability after N-terminal tag cleavage	Purification	Yield (mg/l *E. coli* culture)
	N-terminal	C-terminal				
(334-425)[§]	GST	-	+	degrades	-	-
	GST	His$_6$	+	degrades	+	0.3
	MBP	His$_6$	++	degrades	+	0.7
(340-425)[§]	GST	His$_6$	-	-	-	-
(343-425)[§]	GST	His$_6$	+	degrades	+	1.6
(334-405)	GST	His$_6$	+	stable	+	0.6
	GST	-	+	n.d.	n.d.	
	MBP	His$_6$	++	n.d.	n.d.	
	MBP	-	++	stable	+	0.2
(334-Δ(344-359)-405)	GST	His$_6$	+	stable	+	0.5
	GST	-	+	n.d.	n.d.	
	MBP	His$_6$	++	n.d.	n.d.	
	MBP	-	++	n.d.	n.d.	
(334-Δ(345-359)-405)	GST	His$_6$	+	stable	+	1.0
	GST	-	+	n.d.	n.d.	
	MBP	His$_6$	++	n.d.	n.d.	
	MBP	-	++	n.d.	n.d.	
(334-398)	GST	His$_6$	-	-	-	-
	GST	-	-	-	-	-
	MBP	His$_6$	+	degrades	+	0.9
(334-389)	GST	His$_6$	+	n.d.	n.d.	
	GST	-	-	-	-	-
	MBP	His$_6$	+	stable	+	0.5
(334-382)	GST	His$_6$	+	n.d.	n.d.	
	GST	-	+	n.d.	n.d.	
	MBP	His$_6$	++	stable	+	0.4
(334-375)	GST	His$_6$	-	-	-	-
	GST	-	+	stable	+	0.4
	MBP	His$_6$	++	n.d.	n.d.	
	MBP	-	++	n.d.	n.d.	

Table 3.1: Overview of designed She3p constructs and their expression, stability and purification capabilities. Soluble protein expression of She3p variants, as indicated with C- or/and N-terminal affinity tag(s), was tested in 50 ml small-scale expression tests and was categorized in no (-), good (+) or very good (++) expression. The indicated protein stability refers to She3p protein after cleavage of the N-terminal affinity tag. In case of successful purification (+, ≥ 95 % purity), protein yields are specified. Not determined (n.d.) applies whenever stability properties or purification was not tested. [§] Plasmids expressing these constructs were cloned by Dr. Roland Heym (see Table 2.3).

She3p proteins were initially purified via GST-tag. Several constructs had solubility issues during expression and in some cases resulted in very low protein yields after purification. Both problems could be tempered by purifying MBP-tagged She3p versions. Nevertheless yields were still comparatively low for crystallization trials (Table 3.1). Since degradation occurred in several She3p variants and in case of She3p (334-425) the addition of a C-terminal His_6-tag improved protein stability a C-terminal His_6-tag was added at the C-terminus of each construct. With these measures, degradation could be diminished, but not completely abolished.

Purified C-terminal She3p variants were tested for their ability to form ternary complexes with different versions of She2p and *ASH1*-E3 RNA. Results from analytical size-exclusion chromatography experiments (section 2.10.9) allowed me to categorize the different combinations in three classes (Figure 3.14 and SEC-labeled column in Table 3.2). While the last 20 amino acids of She3p were dispensable for stable ternary complex formation (Figure 3.14 A), the depletion of amino acids 334-342 resulted in a complete loss of She2p-She3p-RNA complexes (Figure 3.14 B). A transient interaction was observed for She3p (334-382)His_6 (Figure 3.14 C). Therefore She3p region 382-405 has probably a rather supportive role in stable ternary complex formation, while amino acids 334-342 seem to be indispensable to bind She2p (6-240, C-S) and E3 (51 nt) in a stable way.
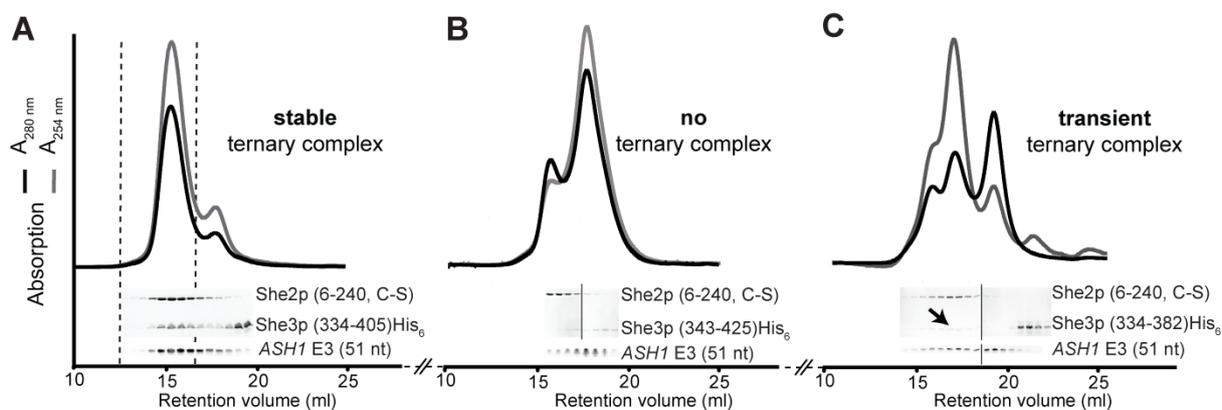
Figure 3.14: Exemplary analytical size-exclusion chromatography runs to test She3p constructs on ternary complex formation. A) She2p (6-240, C-S), *ASH1* E3-(51 nt) RNA and She3p (334-405)His_6 form a stable ternary complex. Fractions, delimited by dashed lines, contained the stoichiometric complex and could be separated from non-stoichiometric content. These fractions were suitable for crystallization. **B)** She3p (343-425)His_6 could not be detected in the fractions containing She2p (6-240, C-S) and E3 (51 nt) and was therefore categorized to be unable to form a ternary complex. **C)** She3p (334-382)His_6 resembles an example for a rather transient complex formation with She2p and the E3 element from A). Arrow marks faint bands that resemble the portion of She3p (334-382)His_6 that co-elutes together with She2p and RNA in a double peak. Complex tests were performed on a Superose 6 10/300 GL column (GE Healthcare). Protein content was analyzed on SDS-PAGE and nucleic acid content on denaturing urea PAGE.

She3p	She2p	ASH1	Synergistic ternary complex formation		Crystals		Remarks
			SEC	EMSA	initial	refined	
(334-425)His$_6$	6-240, C-S	E3 (51 nt)	+	n.d.	+		
(343-425)His$_6$	6-240, C-S	E3 (51 nt)	-	n.d.			Figure 3.14 B
(334-405)	6-240, C-S	E3 (51 nt)	+/-	n.d.	+	+	She2p apo structure, see section 3.3.5.3
(334-405)His$_6$	6-240, C-S	E3 (51 nt)	+	n.d.			Figure 3.14 A
(334-405)His$_6$	6-240, C-S	E3 (28 nt-loop)	+	n.d.			
(334-405)His$_6$	6-246	E3 (28 nt-loop)	+	n.d.	+		
(334-405)His$_6$	6-246, C-S	E3 (28 nt-loop)	+/-	n.d.			
(334-398)His$_6$	6-240, C-S	E3 (51 nt)	+/-	n.d.			
(334-389)His$_6$	6-240, C-S	E3 (51 nt)	+/-	n.d.			
(334-382)His$_6$	6-240, C-S	E3 (51 nt)	+/-	n.d.			Figure 3.14 C
(334-375)	6-240, C-S	E3 (51 nt)	-	yes	+		
(334-Δ(344-359)-405)His$_6$	6-246, C-S	E3 (28 nt-loop)	-	yes	+	+	She2p-She3p structure, see section 3.3.5.4
(334-Δ(345-359)-405)His$_6$	6-246, C-S	E3 (28 nt-loop)	-	yes			

Table 3.2: Overview of She3p-She2p-RNA combinations tested for synergistic ternary complex formation and their crystal yields. Ternary complex formation was tested in analytical size-exclusion chromatography (Figure 3.14) and was categorized to be stable (+), rather transient (+/-) or not formed (-). Qualitative EMSA showed She2p-She3p-RNA binding for some mutants that were not able to form a ternary complex in analytical SEC experiments.

At the beginning She2p-She3p-RNA complexes were separated via size-exclusion to achieve the highest sample homogeneity for crystallization. Complexes composed of different combinations of protein and *ASH1*-E3 constructs were subjected to crystallization trials (Table 3.2 and section 2.12.4). The fact that protein yields of C-terminal She3p fragments were quite low and complex homogeneity was at expense of its yield made crystallization of She2p-She3p-RNA complexes difficult. Apart from several irregular shaped initial crystals that could either not be reproduced or not improved by intensive attempts of refinement, additive screens (section 2.12.5) or seeding (section 2.12.6), two crystallization conditions yielded crystals that were suitable for structure determination.

Firstly, a complex formed by She2p (6-240, C-S), She3p (334-405) and *ASH1*-E3 (51 nt) yielded another She2p apo-structure at 3.9 Å resolution from crystals with a different space group than previously reported (for details see section 3.3.5.3). Here, RNA binding was probably prevented by the high salt concentration in the crystallization condition (1.5 M NH$_4$SO$_4$). As strategic consequence crystals grown in high salt conditions were not followed further in subsequent screenings.

Secondly, the combination of She2p (6-246, C-S), She3p (334-Δ(344-359)-405)His$_6$ and E3-(28 nt-loop) RNA resulted in a She2p-She3p co-structure at 2.80 Å resolution. Apart from the space group and unit cell content, the structure did not largely diverge from the published one [153] (for details see section 3.3.5.4).

It was reported that the addition of trace amounts of proteases to the crystallization drop significantly increases the likelihood of recalcitrant proteins to yield diffraction quality crystals [192]. Regarding the ternary complex, *in situ proteolysis* would degrade flexible, unbound regions of She3p that could hamper crystal formation, while probably not affecting the stably bound parts of She3p. To test this, crystallization drops containing the ternary complex of full-length She3p-His$_6$, She2p (6-246, C-S), and *ASH1*-E3 (28 nt-loop) were supplemented with different proteases. Parameters like protease concentration and incubation temperature were varied (section 2.12.7). The best crystal generated by this method diffracted up to 6.1 Å resolution. Because of the insufficient resolution and the low reproducibility of crystals generated by *in situ* proteolysis [192], this approach was not followed further.

Soaking of small molecules or peptides that penetrate the solvent channels of existing crystals is a common technique to incorporate ligands and to overcome the initial nucleation hurdle. Therefore rare, more voluminous crystals of She2p (6-246, C-S) and the minimal E3 LE were used to soak them (section 2.12.11) with chemically synthesized peptides comprising either She3p region 331-343 (13-mer) or 382-405 (24-mer). In case of the 13-mer peptide the diffraction pattern was usually anisotropic with few spots diffracting up to 3 Å resolution. Soaking the She2p-RNA crystals with the 24-mer peptide improved the diffraction of the co-complex crystals (section 3.3.4.2), but did not provide any insights on how She3p is bound in the ternary complex.

Beneath soaking, also co-crystallization of She2p (6-246, C-S), E3-(28 nt-loop) RNA and either of two different She3p peptides a molar ratio of 2 : 2.1 : 2.5 was tried. Using the 13-mer She3p peptide, plenty of rectangular shaped crystals of ~ 150 μm x 10 μm x 10 μm size were grown at 4 °C after 14 days in 100 mM Bis Tris propane, pH 6.5, 200 mM NaF, 20 % (w/v) PEG3350. Despite their outer beauty they did not show any detectable diffraction. In case of the 24-mer peptide, the crystal shape got worse but the diffraction increased to 8 Å resolution.

Since previous crystallization trials never resulted in a ternary complex structure comprising all three components She2p, She3p and the RNA a new approach with She2p-She3p-fusion proteins was initiated (see section 3.3.5.5).

3.3.5.2 Limited proteolysis

To obtain more information on flexible regions of She3p in the ternary complex and thereby deducing unbound regions of She3p, limited proteolysis combined with Edman degradation was applied (for details see section 2.10.11). The complex of She2p (6-240, C-S), She3p (334-425)His$_6$ and E3 (51 nt) was exposed to five different proteases with varying cleavage properties at different concentrations (Figure 3.15 A). After experimental adjustment, promising protein bands were excised (Figure 3.15 B) and the generated N-termini determined by Edman degradation. Dr. Reinhard Mentele (Max-Planck Institute of Biochemistry, Martinsried, Germany) performed Edman degradation and sequencing.
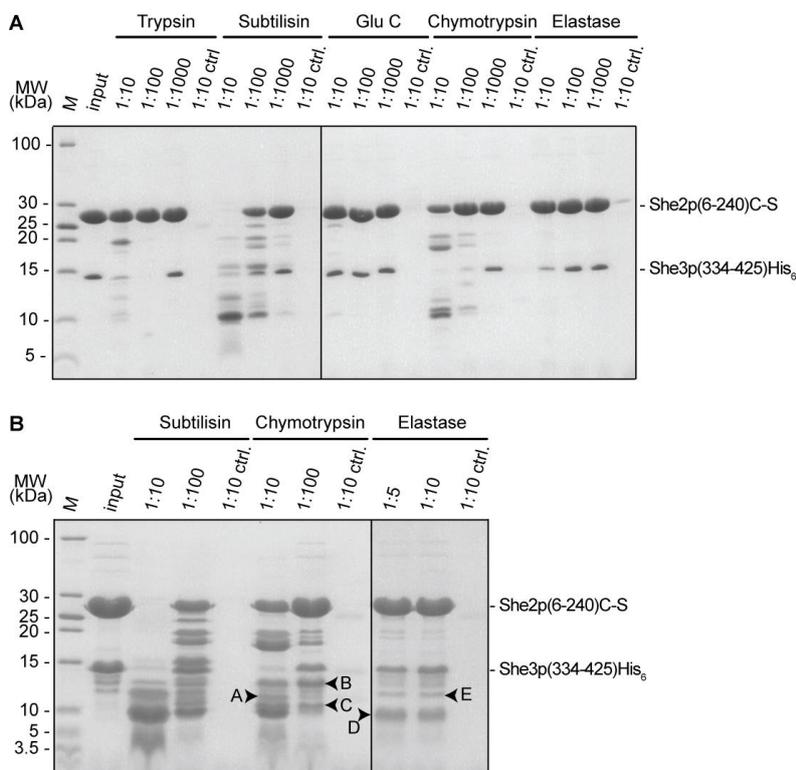
Figure 3.15: Limited proteolysis experiment. The complex of She2p (6-240, C-S), She3p (334-425)His$_6$ and E3 (51 nt) RNA (input) was exposed to different proteases with varying concentrations. Shown are 17 % SDS PAGE gels stained with PageBlue. **A)** Initial optimization experiment showing degradation products of She2p and She3p generated by trypsin, subtilisin, Glu C protease, chymotrypsin or elastase at different protease concentrations (1:10, 1:100, 1:1000 (w/w)). Proteases alone were used as control (ctrl., dilution 1:10). **B)** According to the results from A), promising proteases subtilisin, chymotrypsin and elastase were subjected to a higher complex content and the protease concentrations were adjusted. Arrows mark bands (A-E) that were excised and prepared for Edman degradation.

Sequencing results showed that samples deriving from bands 'A' and 'C-E' (Figure 3.15 B) just revealed the very N-terminus of She3p (334-425)His$_6$. Sole band 'B' consisted of She3p N-termini with the sequences ^{346}TASPLL^{351} and ^{345}YTASPL^{350}, detected at a 1 : 1 ratio. Thus, amino acids in this area seem to be flexible in the She2p-She3p-RNA complex. Considering these results and the She3p sequence alignment from different yeast strains (Figure S13 in reference [139]), two more She3p mutants (334-Δ(344-359)-405)His$_6$ and (334-Δ(345-359)-405)His$_6$ (Table 3.1) were added to the crystallization trials for the ternary complex.

3.3.5.3 She2p apo structure – Space group P2$_1$2$_1$2

A sample of She2p (6-240, C-S), She3p (334-405) and *ASH1*-E3 (51 nt) was prepared by mixing the correct stoichiometric ratio of She2p and RNA and adding a molar excess of She3p. Different complex species were separated on a Superose 6 10/300 GL column (GE Healthcare) in 20 mM Hepes/NaOH, pH 7.8, 200 mM NaCl. According to SDS-PAGE and denaturing PAGE, fractions containing complexes in the correct stoichiometric ratio were concentrated to 2.24 mg/ml. Initial screens were performed at 21 °C at the MPIB Crystallization Facility (Max Planck Institute of Biochemistry, Martinsried, Germany) and yielded star-fruit shaped crystals in 50 mM Na cacodylate pH 5.5, 1.7 M NH$_4$SO$_4$ and 5 mM MgOAc (Figure 3.16 A). Some of the refined crystals growing in 50 mM Na cacodylate pH 5.3, 1.5 M NH$_4$SO$_4$, 5 mM MgOAc, 4 % (w/v) adonitol after 10 days, 21 °C, using the hanging drop vapor diffusion method with 200 µl reservoir solution and 2 µl drop size, changed their shape to single pyramids (Figure 3.16 B). Crystals were cryogenically protected by incubating them for 3 s in 50 mM Na cacodylate pH 5.3, 1.5 M NH$_4$SO$_4$, 5 mM MgOAc, 400 mM trimethylamine/Cl, 25 % (v/v) ethylene glycol and flash-frozen in liquid nitrogen.

Data collection at 100 K was performed at beam line P11 at the Deutsches Elektronen Synchrotron (DESY, Hamburg, Germany). Data were indexed, integrated and scaled using the XDS program package [204]. Structure factor amplitudes were obtained via the Truncate procedure [195]. Initial electron-density maps were calculated by molecular replacement with the apo-structure of She2p (PDB-ID: 1XLY) as search model using the program Phaser [205]. Data collection and processing details are summarized in Appendix Table 5.3. Due to the rather low resolution of 3.9 Å, and the fact that neither RNA nor She3p was visible, the structure was just once refined with RefMac 5.7 [210, 211]. Compared to the published She2p apo-structure [150] the space group changed from

monoclinic (*C*2) to orthorhombic (*P*2$_1$2$_1$2) and instead of two four She2p molecules were present per asymmetric unit (Figure 3.16 C).
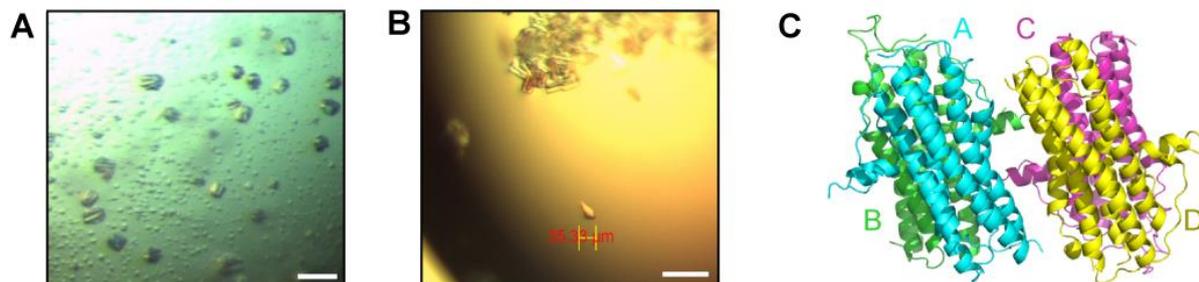
Figure 3.16: Crystals grown from She2p (6-240, C-S), She3p (334-405) and *ASH1* E3 (51 nt) yielding a She2p apo-structure. A) Star-fruit shaped crystals grown in 50 mM Na cacodylate pH 5.5, 1.7 M NH$_4$SO$_4$ and 5 mM MgOAc 50 mM Na cacodylate pH 5.5, 1.7 M NH$_4$SO$_4$ and 5 mM MgOAc. c$_{complex}$ = 2.24 mg/ml. **B)** Crystals grown in 50 mM Na cacodylate pH 5.3, 1.5 M NH$_4$SO$_4$, 5 mM MgOAc, 4 % (w/v) adonitol. Rarely single pyramidal crystals appeared. Depicted scale bar in **A)** and **B)** represents 100 µm. **C)** Crystal structure of She2p at 3.9 Å. Each of the four She2p monomers is colored according to their chain IDs (A-D). Chain A and B as well as chain C and D form homodimers, which is in accordance to the published She2p structure [150].

3.3.5.4 Structure of She2p and a She3p peptide – Space group P2$_1$

Samples containing 92.4 µM She2p (6-246, C-S), 106.3 µM She3p (334-Δ(344-359-405)-His$_6$ and 46.2 µM *ASH1* E3 (28 nt-loop) RNA in 20 mM Hepes/NaOH pH7.8, 200 mM NaCl, 2 mM MgCl$_2$ were incubated at room temperature for 10 min and spun down for 20 min, 16,100 x g, 4 °C. 1 µl of the protein/RNA-containing solution was mixed with an equal amount of well solution (50 mM Hepes pH 7.0, 45 % (v/v) tacsimate pH 7.0, 2 mM spermine, 2 mM hexamine cobalt(III)chloride). Crystals, grown at room temperature using the hanging drop vapor diffusion technique, were frozen in mother liquor supplemented with 25 % (v/v) ethylene glycol.

Native diffraction data of one single crystal was recorded at 100 K at beam line ID23-2 at the ESRF (Grenoble, France). Data processing was done as described for She2p in section 3.3.5.3. Phases were improved by using non-crystallographic symmetry (NCS) averaging during refinement with RefMac5.8 [210, 211]. The structure of She2p and parts of She3p at 2.8 Å resolution was completed by manual building in the program COOT [206]. Data collection and refinement statistics are listed in Appendix Table 5.3

The crystal structure revealed in total eight She2p monomers, each bound by one She3p peptide per asymmetric unit cell (Figure 3.17). RNA was again missing in the structure due to neighboring She2p molecules in the crystal lattice that blocked part of the RNA binding region. Comparing this structure to the published She2p-She3p co-structure [153], the crystal space group changed from monoclinic *C*2 [153] to *P*2$_1$ and the unit cell content increased. Overall C$_\alpha$ atoms show no major conformational changes compared to its equivalents in the published co-structure (RMSD of PDB-ID 4WNL = 0.70 Å). While two She2p homodimers (A, B and C, D) form the same tetramer as in the She2p-RNA co-complex structure (section 3.3.4.3), the third and forth set of homodimers (E, F and G, H) form tetramers with the contents of the neighboring unit cell. Regarding She3p (chains I-P) at least the conserved LPGVK motif (positions 364-368) [153] is visible in each of the eight chains. The longest She3p stretch that could be modeled was chain J and comprised ten amino acids (SPVLPGVKRT reflecting positions 362-369). Serine 362 and threonine 369 were the only residues that were new compared to the published structure [153].
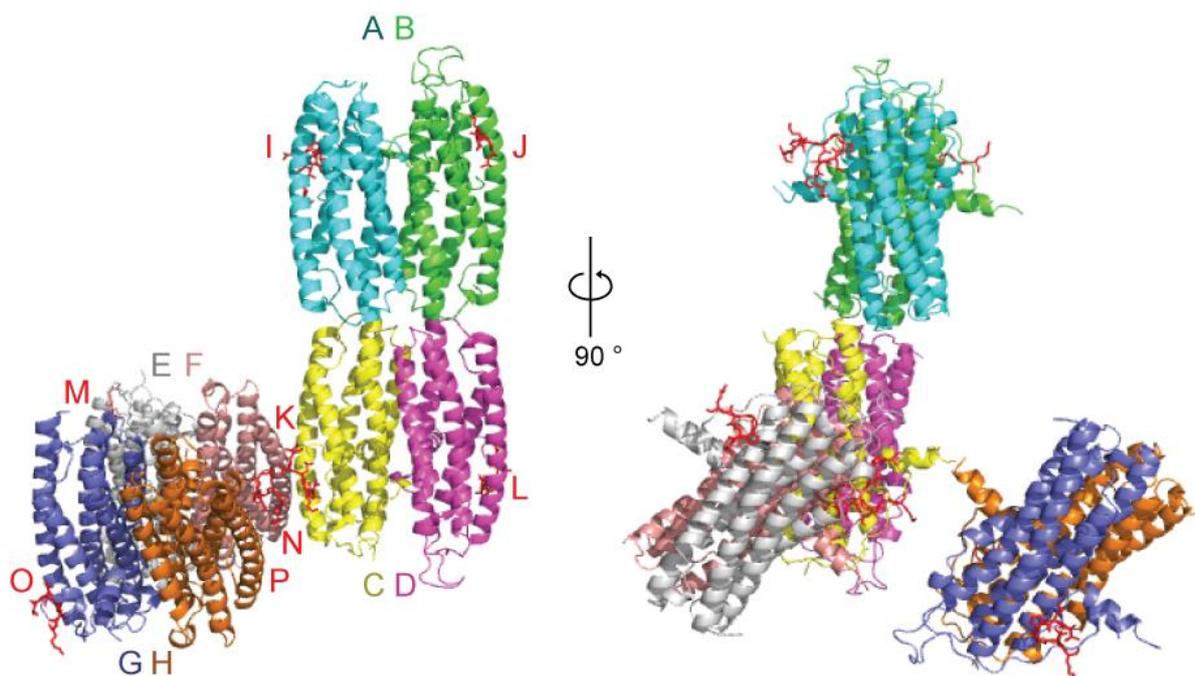
Figure 3.17 Crystal structure of She2p and parts of She3p. Crystallizing She2p (6-246, C-S), She3p (334-Δ(344-359)-405)-His$_6$ and *ASH1* E3 (28 nt-loop) RNA revealed just She2p and parts of She3p being present in the structure. The content of one unit cell is shown from the front (left) and side view (vertically turned around 90 °, left). Each of the eight She2p monomers is colored according to their chain IDs (A-H). The eight visible She3p fragments (chain IDs I-P) are colored in red.

3.3.5.5 She2p-She3p fusion constructs

In previous crystallization trials She3p was never present in the crystal together with She2p and the RNA. To increase the likelihood of She3p incorporation in the crystals and to enforce complex formation in the correct stoichiometric ratio, She2p and She3p were covalently connected via a flexible linker. *SHE2(6-246, C-S)* was C-terminally fused via one or two *(GGSGG)$_{1/2}$* linker(s) to a range of *SHE3* constructs (FE#30-#52 in Table 2.3). Expression tests in different *E. coli* strains revealed most efficient protein production in *E. coli* BL21 Gold (DE3) pRARE. Although bacterial growth in auto-induction medium resulted in higher total protein expression, using LB medium instead yielded comparatively more soluble protein. Several purification strategies including N-His$_6$-MBP and N-His$_6$-SUMO tag were tried. Although final yields of protein were small (purity > 95 %), N-His$_6$-GST-tag worked out best for most constructs (Figure 3.18). Established purification protocols are described in detail in section 2.9.9.

A

Region of She3p fused to She2p (6-246, C-S)	N-terminal affinity tag	No. of flexible GGSGG linker	Yield (mg/ l E.coli culture)	Corresp. lane No. in B)
- (control)	His$_6$-GST	2	4.54	-
(331-405)	His$_6$-GST	1	0.08	6
	His$_6$-GST	2	0.19	7
	His$_6$-MBP	2	-	-
(331-375)	His$_6$-GST	1	0.06	4
	His$_6$-GST	2	0.04	5
	His$_6$-MBP	1	-	-
	His$_6$-SUMO	1	-	-
(331-343)	His$_6$-GST	2	0.12	3
	His$_6$-MBP	1	4.95	-
(331-405)Δ(345-359)	His$_6$-GST	1	0.18	8
	His$_6$-GST	2	0.08	9
	His$_6$-MBP	1	-	-

B
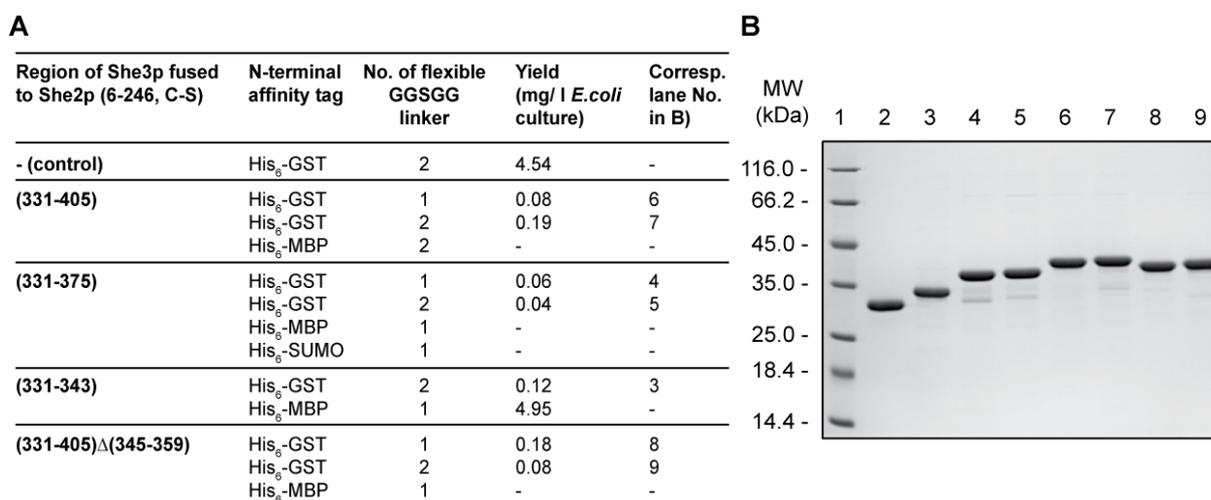
Figure 3.18 Overview of She2p-She3p-fusion constructs. A) Table summarizing a range of She2p(6-246, C-S)-(GGSGG)$_{1/2}$-She3p-fusion constructs that were tested for purification. N-terminal affinity tags were cleaved in course of purification. "- (control)" indicates the control protein, which was just She2p(6-246, C-S) fused to the long (GGSGG)$_2$ linker. Given yields correspond to proteins after the final purification step. (-) indicates problems during purification: His$_6$-SUMO tag cleavage was not efficient and could not be satisfyingly improved. His$_6$-MBP tags could not be sufficiently separated from the protein of interest. **B)** Quality control of purified proteins after freezing and thawing shows stability and purity of > 95 % in each case. Lane numbers correspond to She2p-She3p fusion proteins as indicated in the table of A). Apart from that, lane 1 reflects the molecular weight marker and lane 2 She2p (6-246, C-S). Molecular weights are assigned as follows: 2 = 28.0 kDa, 3 = 29.8 kDa, 4 =32.7 kDa, 5 = 33.1 kDa, 6 = 35. 8 kDa, 7 = 36.1 kDa, 8 = 34.3 kDa and 9 = 34.6 kDa. 3 µg of each protein was loaded on a 12 % SDS-PAGE gel and stained with PageBlue.

Purified She2p-She3p fusion proteins were tested for *ASH1*-E3 (28 nt-loop)-binding in EMSAs (Figure 3.19). In general all fusion proteins showed high affinity ternary-complex formation and linker

lengths had just minor effects on RNA-binding affinities. Shortening the She3p-fragment length from the C-terminus lead to a continuous decrease in synergistic binding affinity. Nevertheless the shortest fusion construct (She2p(6-246, C-S)-(GGSGG)$_1$-She3p(331-343)) still had an estimated K_D of ~ 600 nM (Figure 3.19 B, second last gel) and was therefore also appropriate for initial crystallization screening.

Figure 3.19: EMSAs to test She2p-She3p linker constructs. A) Schematic representation of protein components in linker constructs. She2p (6-246, C-S) was C-terminally fused via one or two flexible GGSGG linkers to different She3p variants (dark green). Light green areas

indicate UV-crosslinking regions in She3p with E3 (51 nt) ([139] and unpublished results by Dr. Roland Heym). **B)** EMSAs with radioactively labeled *ASH1*-E3 (28 nt-loop)-RNA and She2p(6-246, C-S)-(GGSGG)$_{1/2}$-She3p linker constructs show complex formation for each construct. She2p(6-246, C-S)-(GGSGG)$_2$ alone, termed "S2-LL" served as control and did not give any band shift. Increasing protein concentrations were 20 nM, 60 nM, 180 nM, 540 nM, 1.61 µM, 4.86 µM. Experiments were performed in triplicates.

3.3.5.6 Crystallization trials with She2p-She3p fusion constructs and RNA

Initial crystallization screens using the shortest fusion construct She2p(6-246, C-S)-(GGSGG)$_1$-She3p(331-343) together with *ASH1*-E3 (28 nt-loop) yielded few well-shaped crystals in two conditions. Best crystals, grown in 100 mM Bis Tris propane, pH 6.5, 20 % (w/v) PEG 3350, 200 mM NaBr, at 4 °C, yielded an anisotropic diffraction pattern, but also showed some spots up to 2.6 Å resolution. Crystals could be rarely reproduced and did not improve when using an additive or heavy atom screen (section 2.12.5). They looked even worse in refinement experiments, although different protein, salt and precipitant concentrations as well as different pHs and varying sodium counter ions were tested. Nevertheless, when testing an initial crystal at the synchrotron a data set could be collected. After processing and molecular replacement with the co-complex of She2p and the RNA (section 3.3.4.3), eight She2p monomers and four RNA molecules were visible per unit cell. The crystal had the same *C2* symmetry and a slightly worse resolution of 2.95 Å. Unfortunately no additional density deriving from the She3p part of the fusion protein could be detected.

The variant She2p(6-246, C-S)-(GGSGG)$_1$-She3p(331-405) Δ (345-359) together with the minimal E3 LE yielded tiny crystals that did not diffract and were therefore not followed further.

Apart from the shortest She2p-She3p fusion protein, the most promising variant was the longest fusion construct She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405) with *ASH1*-E3 (28 nt-loop). Here, initial crystals grew in 200 mM NaI, 20 % (v/v) PEG 3350 at 4 °C and yielded a data-set with 4.0 Å resolution that finally revealed parts of She3p present in the She2p:RNA structure. On the one hand, few more amino acids additional to the recently published She3p peptide (She3p (362-369)) [153] could be modeled. On the other hand 12 extra amino acids were placed into an extra stretch of electron density close to C1779 and U1780. Since the resolution of the data set was just 4 Å, those amino acids could not be assigned to a specific She3p position. Nevertheless due to the distance of those residues with respect to the C-terminus of She2p and the known position of She3p (362-369) some could potentially belong to the region She3p (334-340), which was shown to UV-crosslink to RNA

[139]. Since one part of the additional She3p density was in proximity to the 3' end of *ASH1*-E3 (28 nt-loop)-RNA, the idea to extend the 3' end by several bases was followed.

In this respect four more RNAs with varying 3' ends were obtained from commercial suppliers. Crystallization of She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405) together with the extended RNAs E3-A (29 nt), -B (29 nt) and -C (30 nt) (Table 2.5) did not yield any crystals. The combination of She2p(6-246, C-S)-(GGSGG)$_1$-She3p(331-343) and the extended E3-D (31 nt) RNA (Table 2.5) resulted in crystals with the shape of thick needles. Although crystals grew in different conditions and were extensively refined, none of them showed any detectable diffraction.

Going back to the best diffracting crystals of She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405) and the E3 (28 nt-loop) RNA more extensive refinement was pursued. Parameters like complex concentration, drop size, type of sodium counter ion, PEG concentration, pH and ratio of protein/RNA complex versus well solution were varied. Thin and plate-shaped crystals still grew seldom and slow. This fact did not change when I tried streak or micro seeding (section 2.12.6). Also changing crystallization trials from the hanging drop to the sitting drop vapor diffusion method had no effect on crystal's appearance. Subjecting the complex to an additive, heavy atom and detergent screen (section 2.12.5) yielded few, but more irregular shaped crystals.

Since it was reported that diffraction quality of She2p-She3p crystals improved considerably after chemical crosslinking [153], different crosslinking approaches with varying crosslinking reagents, concentrations, incubation times and application methods were tested (for details see section 2.12.9). None of these combinations improved diffraction quality beyond 7 Å resolution.

Crystallization using the free interface diffusion technique [215] was reported to produce few, but large crystals. I applied this approach also to my crystallization trials and tested capillary crystallization with varying parameters (for details see section 2.12.10). The tested set-ups did not yield any crystals.

Another idea was to collect experimental phase information in the region of the free electron density to assign the missing She3p amino acids. Since this method did not give the desired insights (for details see the following section 3.3.5.7) a systematic alanine screen of the assumed binding region of She3p was performed (for details see the following section 3.3.5.8). Here, EMSAs revealed that residues K340, R341 and R342 of She3p are involved in E3-LE binding. Trials to fill electron densities

in proximity to the RNA by shifting this 'KRR'-pattern together with their surrounding amino acids did not improve R_{work} and R_{free} values.

Low diffracting plate-like crystals had not just the problem of being thin, but also big issues by means of reproducibility. To overcome these problems crystallization stock solutions were directly purchased from the same manufacturer that also provided the initial screening libraries. Fortunately the reproduction of the crystals diffracting to 4 Å resolution improved remarkably by using these bought stock solutions.

Subsequent crystal refinement started from the beginning. The precipitant concentration of PEG 3350 was optimized to 21 %, which yielded slightly more crystals. While seeding (section 2.12.6) as well as an additive and detergent screen did not improve crystal's appearance, varying the salt concentration and the sodium counter ion ($NaNO_3$, NaI, Na/K tartrate, NaF and NaBr) finally yielded a crystallization condition with crystals diffracting to 2.8 Å resolution. With data sets obtained from these crystals the structure of the mature, cytoplasmic mRNA-recognition complex could be unambiguously solved (section 3.3.5.10). Crystallization and structure determination details are described in detail in section 3.3.5.9.

3.3.5.7 Selenomethionine She2p-She3p fusion constructs

The She3p sequence could not be assigned to the electron density in the best diffracting crystal. To overcome this problem experimental phase information should be generated for individual residues, potentially allowing for an unambiguous assignment of the sequence register to the electron density. Because She3p does not have any methionines in relevant positions, three different point mutants of the She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405) construct, N333M, K340M and S343M were generated (FE#62-#64 in Table 2.3). SeMet substituted proteins were expressed in bacteria (section 2.9.1) and purified as described in section 2.9.9. EMSAs with E3 (28 nt-loop) RNA showed that selenomethionine incorporation did not severely affect complex formation (Figure 3.20). Mutants were therefore used in initial crystallization screening (section 2.12.4).
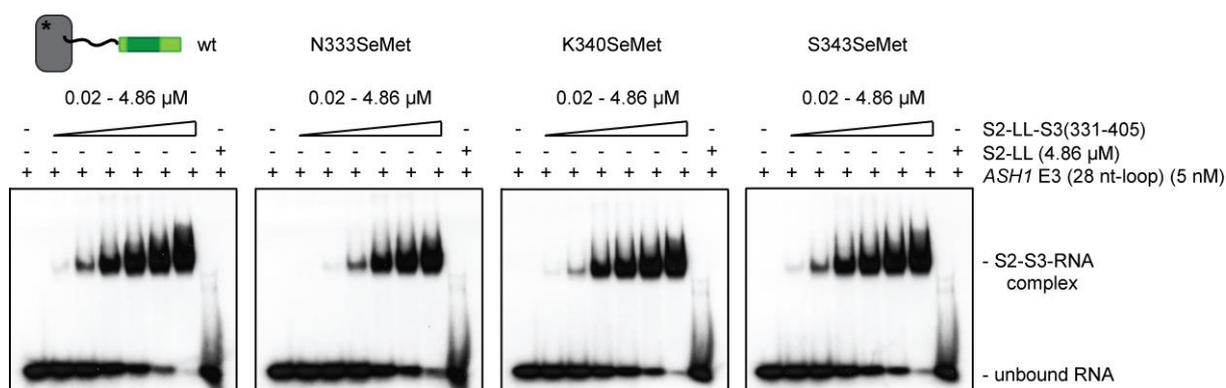
Figure 3.20: EMSAs with selenomethionine She2p-She3p linker constructs. Complex formation of She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405) ("wt", "S2-LL-S3(331-405)") with radioactively labeled *ASH1*-E3 (28 nt-loop)-RNA was compared with three different SeMet mutants (N333SeMet, K340SeMet and S343SeMet). She2p (6-246, C-S)-(GGSGG)$_2$ alone, termed "S2-LL" served as control and showed no complex formation. Experiments were performed in triplicates.

Initial screening of the three She2p-She3p-N333SeMet, -K340SeMet and -S343SeMet mutants together with the minimal E3-LE yielded several crystals in different conditions. Nevertheless, crystals grew again slower and appeared after ~ 20 d. After refining the best conditions by varying sodium counter ions, complex concentration and precipitant concentration plate-shaped crystals were obtained and tested at the synchrotron. A fluorescence scan verified the presence of selenium in the crystals. While crystals including either the N333SeMet or K340SeMet fusion protein showed only low-resolution diffraction patterns (> 15 Å), crystals containing mutant S343SeMet, grown from 100 mM Bis Tris propane, pH 8.3, 20 % (w/v) PEG 3350, 200 mM Na/K tartrate yielded a data set, which diffracted to 4.3 Å resolution. Nevertheless, the anomalous scattering of selenium atoms could not be detected in the electron density and therefore no She3p sequence was assigned.

3.3.5.8 Alanine screen in She2p-She3p fusion constructs

To better understand which amino acids in She3p might contribute to complex formation and thus would potentially be located within the visible electron density close to the RNA of the ternary complex (section 3.3.5.6), a systematic alanine screen in She3p was performed. Several selected alanine mutations were introduced to the She3p part of She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-343) (FE#65-#67 and FE#72-#75 in Table 2.3, Figure 3.21 A). Alanine fusion mutants were purified as described in section 2.9.9 and their binding to E3-(28 nt-loop) RNA assessed in EMSA (Figure 3.21 B). In contrast to single amino-acid mutations N333A, N335A, N338A and N339A that showed wild-type RNA-binding, each of the mutations K340A, R341A and R342A revealed strongly reduced E3-LE binding (Figure 3.21 B). Consequently, the combination of the latter mutations in She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-343)(K340A, R341A, R342A) lead to abrogated ternary complex formation (Figure 3.21 C). Since the tested fusion constructs just carried 13 amino acids of She3p, the impact of mutations K340A, R341A and R342A was also assessed in a longer context. Therefore alanine mutants of She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405) (FE#77-#80 in Table 2.3) were produced and also assessed in EMSA experiments (Appendix Figure 5.8, Figure 3.21 D). Here RNA-binding defects were by far milder but still detectable. Together these data indicate that at least parts of She3p region 344-405 also mediate E3 (28 nt-loop) binding, while She3p residues K340, R341 and R342 could be directly implicated in the process of RNA binding.
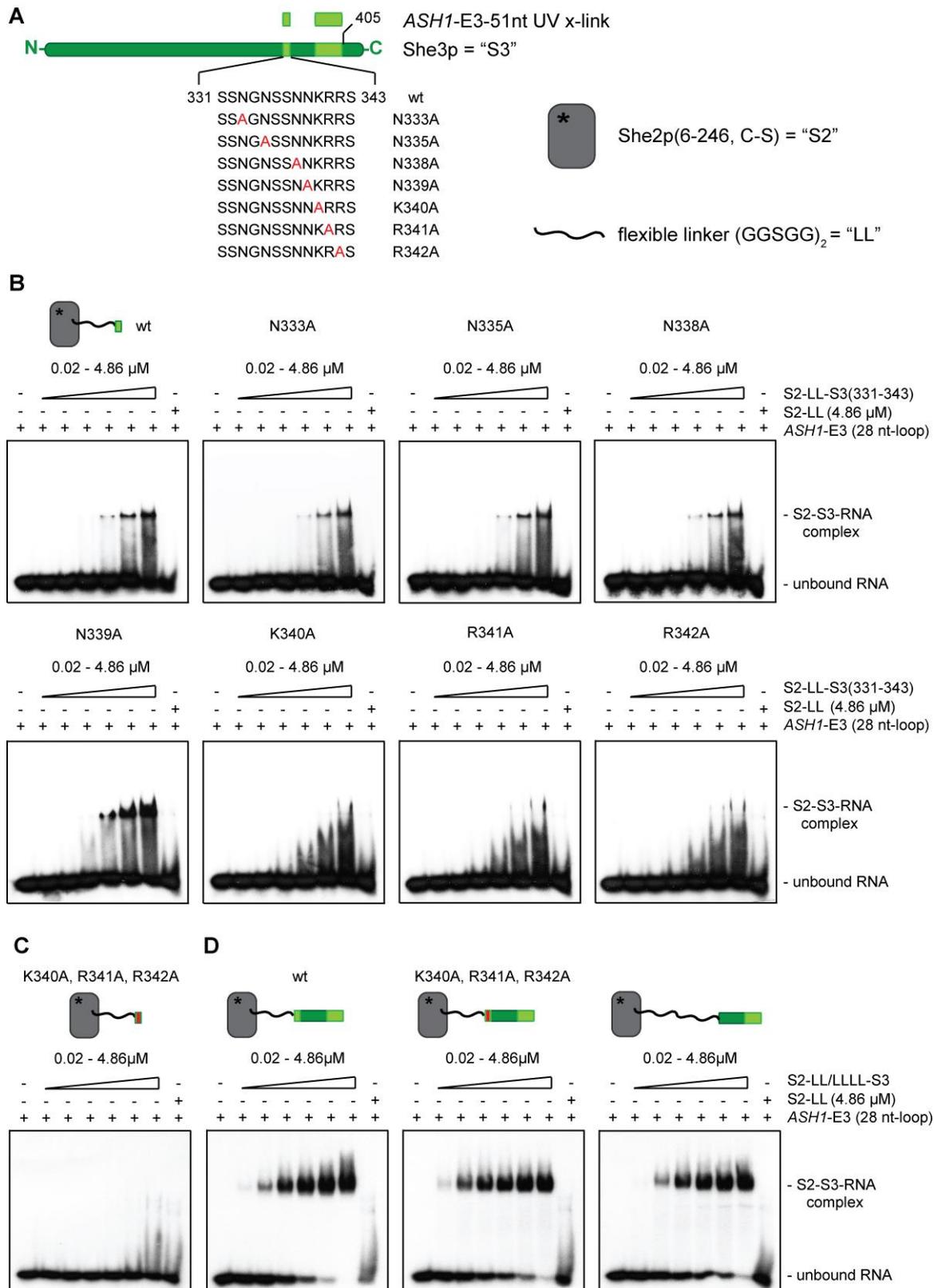
Figure 3.21 Assessment of She2p-She3p alanine mutants. A) Schematic representation of full-length She3p ("S3") and sequential alanine mutation's in the fused She3p region 331-343 ("wt"). Seven mutations including N333A, N335A, N338A, N339A, K340A, R341A and R342A

are tested in B). **B)-D)** EMSAs testing different She2p(6-246, C-S)-(GGSGG)$_2$-She3p fusion proteins ("S2-LL-S3") for their E3 (28 nt-loop) RNA-binding ability. Increasing protein concentrations are 20 nM, 60 nM, 180 nM, 540 nM, 1.62 µM and 4.86 µM. She2p(6-246, C-S)-(GGSGG)$_2$ termed "S2-LL" together with the minimal E3-LE served as control and does not show any band shift. **B)** The RNA-binding ability of the fusion protein containing She3p residues (331-343) termed "wt" is compared to its alanine mutants depicted in A). Just mutants K340A, R341A and R342A lead to impaired RNA binding. **C)** She2p-She3p(331-343)(K340A, R341A, R342A) triple mutant showing abolished RNA binding. **D)** The She2p-She3p(331-405) fusion termed "wt" is compared to its K340A, R341A, R342A triple mutant and a construct where the N-terminal 13 amino acids of She3p are displaced by (GGSGG)$_2$-GGS linker (She2p(6-246, C-S)-(GGSGG)$_4$-GGS-She3p(344-405)). The latter (gel in the middle and right) show slightly diminished ternary complex forming ability. All EMSAs were performed in triplicates.

3.3.5.9 Crystallization and structure determination of the She2p-She3p-ASH1 E3 complex

The complex was assembled using 2.6 mg/ml She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405) and a slight stoichiometric excess of RNA over protein (n[She2p-(GGSGG)$_2$-She3p] : n[E3 (28 nt-loop)] = 2:1.2), in a final buffer of 20 mM Hepes/NaOH pH 7.8, 200 mM NaCl and 2 mM MgCl$_2$. After centrifugation at 16,100 x g, for 30 min at 4 °C, 1 µl co-complex was mixed with 1 µl well solution using the hanging-drop vapor diffusion method in 24-well VDX plates (Hampton Research) with a reservoir volume of 500 µl. Plate-shaped crystals grew at 4 °C in 21 % (w/v) PEG3350 (Hampton Research), 150 mM NaNO$_3$ after approximately two weeks. They were cryogenically protected by 20 % (w/v) PEG3350, 200 mM NaNO$_3$, 25 % (w/v) ethylene glycol and flash-frozen in liquid nitrogen.

Native X-ray diffraction data was collected at 100 K at the X06DA (PXIII) beam line at Swiss Light Source (SLS, Villigen, Switzerland). Processing was done as described for the She2p/RNA co-complex. The previously published apo-structure of She2p (PDB-ID 1XLY) was then also taken for molecular replacement using MolRep [209]. Further model building and refinement with the native data set (2.80 Å) was performed with the programs COOT [206] and RefMac5 [210, 211] including NCS averaging. Data collection, processing and refinement parameters are listed in Appendix Table 5.2. The final structure is deposited at the protein data bank (PDB-ID: 5M0J).

3.3.5.10 Crystal structure of the mature cytoplasmic mRNA-recognition complex

The ternary structure was solved and refined as described in section 3.3.5.10. One unit cell of the crystal lattice consists of four She2p (6-246, C-S), two E3 (28 nt-loop) RNA and four She3p molecules. Although the latter were covalently connected to the She2p C-termini in the crystallized complex, the flexible (GGSGG)$_2$-linkers were not visible (Figure 3.22 A).
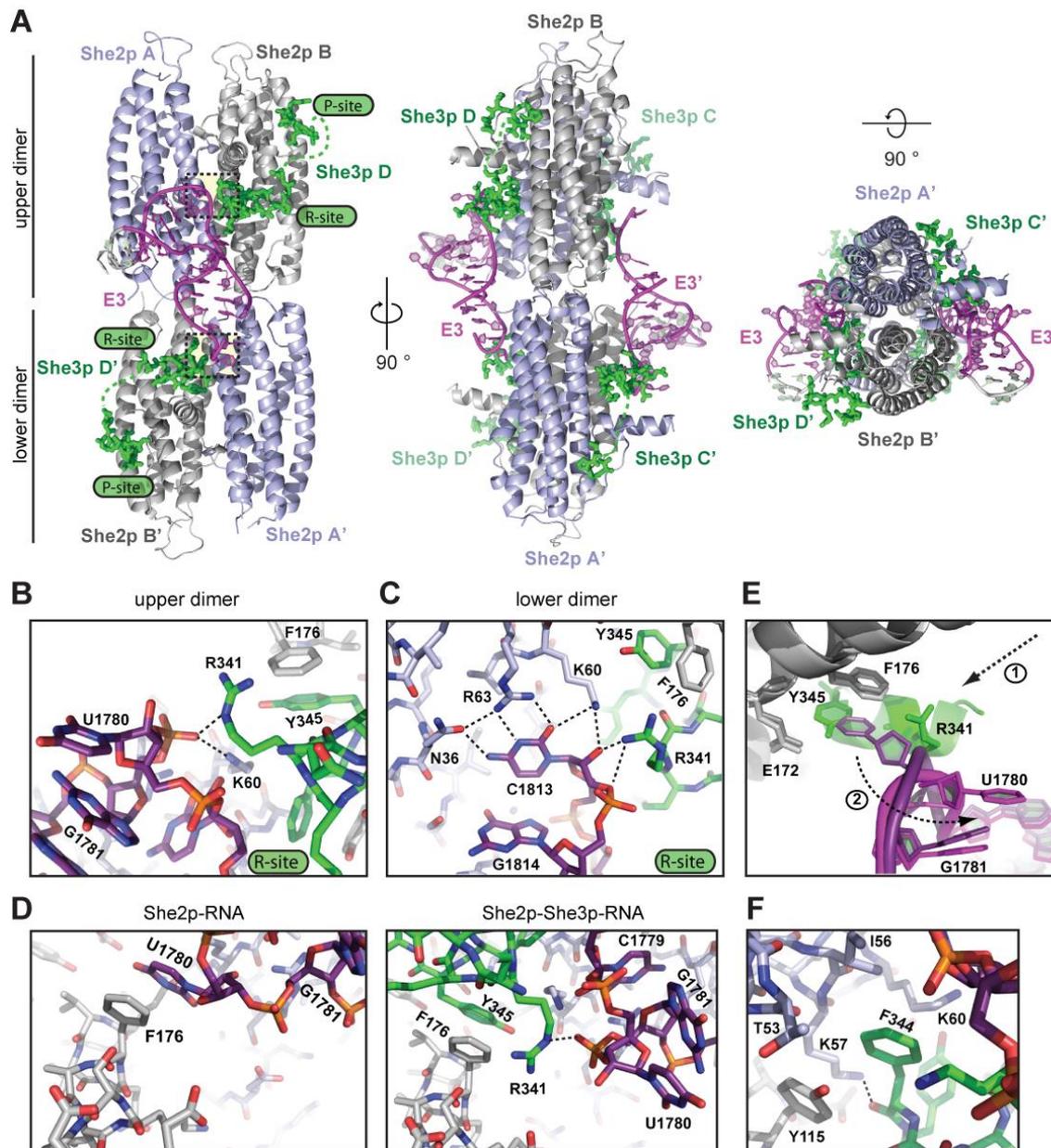
Figure 3.22: Structure of the synergistic ternary complex of She2p, She3p and the minimal *ASH1* E3-localization element. A) Crystal structure of She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405) and the E3 (28 nt-loop) RNA at 2.80 Å as cartoon (She2p/RNA) and stick (She3p) representation. The flexible linker was not visible. Color-coding is equivalent to **Figure 3.10** and visible parts of She3p are depicted in green. The four chains are labeled with She3p C, C', D and D'. Boxed regions in the upper and lower dimer are shown as close-ups in B) and C), respectively. **B)** Close-up of the She3p R-site contact to E3 in the upper dimer. R341 of She3p and K60 of She2p form stabilizing salt-bridges to the phosphate of U1780, while Y345 of She3p base stacks with F176 of the protruding helix in She2p. **C)** Close-up of the She3p R-site contact to E3 in the lower dimer. Here R341 of She3p contacts the phosphate of the conserved C1813 at the 3' end of the RNA. As shown in B) Y345 also interacts with F176 in the lower dimer. The coordination of C1813 by She2p is additionally supported by K60 in the ternary complex. **D)** Comparison of the local environment of She2p F176 in the binary (left) and the ternary complex (right). While F176 of the protruding helix base stacks with U1780 in the co-complex, this interaction is displaced by the Y345-F176 interaction upon joining of She3p. In consequence U1780 rotates away and gets solvent exposed. One of its phosphates is additionally stabilized by R341. **E)** Overlay of close-

ups from U1780 in the co-complex (She2p in gray, RNA in pink) and the ternary complex (She2p in light-gray, RNA in light-pink, She3p in green). (1) Upon squeezing of the She3p R-site in between the protruding helix of She2p and the E3 LE, U1780 is displaced by the Y345 from its interaction with F176 and E172. (2) In consequence U1780 flips away to make space for She3p. **F)** The conserved F344 in the R-site of She3p is highly coordinated by several hydrophobic interactions to She2p (T53, I56, K60, and Y115) and a salt bridge with K57 additionally stabilizes its carbonyl oxygen.

For each She3p molecule two different electron density stretches could be assigned (Figure 3.22 A). The first fragment contained residues 361-367 and overlapped partially with the published She3p peptide from reference [153]. Since this part of She3p makes exclusively protein-protein contacts with She2p it was termed "P-site". The second fragment could be assigned to She3p residues 338-351 and binds to both RNA and She2p and is therefore called "R-site". While the latter buries a surface area of 922 $Å^2$ the P-site masks a smaller region of 630 $Å^2$.

The R-sites of She3p intercalate between the E3 LEs and the protruding helices of each She2p monomer and form in three out of four cases small α-helices. Here She3p R341 plays an important role since it recognizes the E3 RNA together with K60 of She2p at two different positions. Both amino acids contact on the one hand the phosphate backbone of U1780 residing in the kinked region of the RNA (Figure 3.22 A, B) and on the other hand C1813 at the 3' end of the RNA (Figure 3.22 A, C). In addition to R341 of She3p also its conserved K340 supports RNA binding by contacting the phosphate of A1778 in the RNA (Figure 3.23).
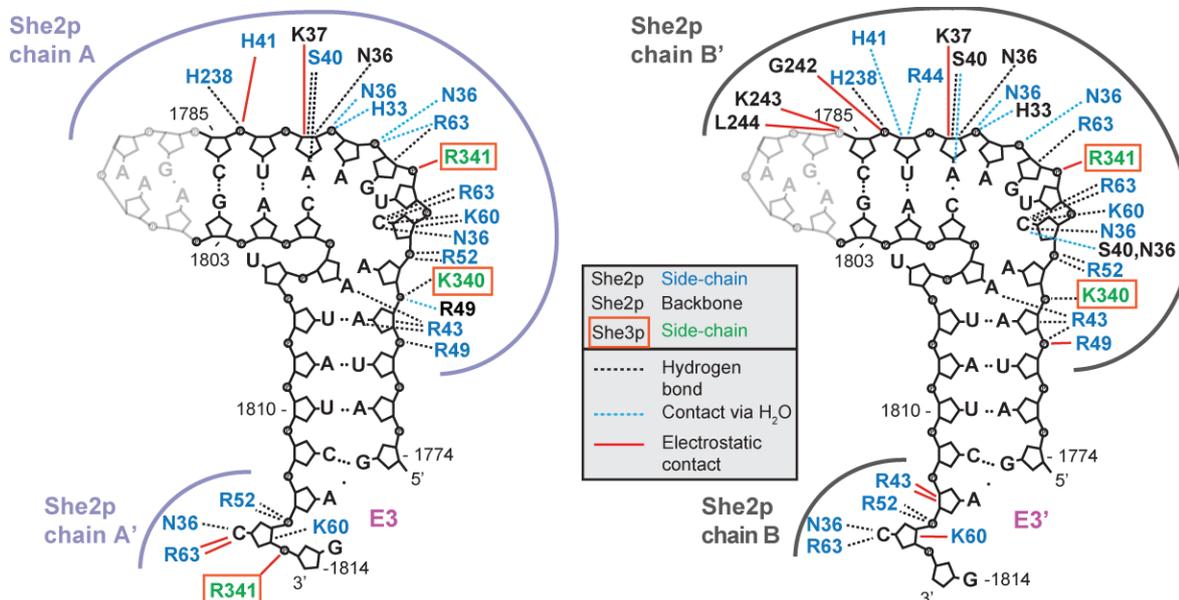
Figure 3.23: Schematic representation of *ASH1* E3-RNA contacts to She2p and She3p. Color-coding of the different chains is equal to **Figure 3.22**. Protein-RNA contacts derive from Nucplot [196] analysis. She3p contributes with the side-chains of K340 and R341 to phosphate-backbone recognition of E3 (28 nt-loop) RNA.

Comparing the quaternary recognition of the conserved cytosines C1779 and C1813 in the She2p-RNA co-complex with the recognition in the ternary complex, the interaction is maintained and additionally strengthened via K60 of She2p and R341 of She3p (Figure 3.22 C and Figure 3.23). In contrast, the third sequence specific interaction of U1780 that base stacks with F176 and interacts with E172 of She2p in the co-complex becomes released upon joining of She3p in the ternary complex (Figure 3.22 D and Figure 3.11 A, D). As consequence U1780 rotates away, becomes solvent exposed and Y345 of She3p base stacks with F176 of She2p instead (Figure 3.22 E). This F176-Y345 interaction can be observed in the protruding helices of all four protomers of the She2p tetramer.

As already mentioned the R-site does not solely make RNA-contacts. One of the highest coordinated protein-protein interactions between the She3p R-site and She2p is represented by the conserved F344 (Figure 3.22 F). It has several hydrophobic interactions with T53, I56, K60, and Y115 of She2p and its carbonyl oxygen is additionally stabilized via a salt bridge to K57. The numerous She2p-She3p R-/P-site protein-protein contacts are summarized in Appendix Table 5.4 and more explicitly shown in Appendix Figure 5.12 and Appendix Figure 5.13. Also intramolecular protein-protein interactions play important roles in complex stabilization. She3p R342 in the R-site stabilizes the She3p R-site helix by side-chain contacts to the carbonyl oxygens of S348 and P349 of the same peptide chain (not shown).

3.3.6 Structure validation

3.3.6.1 Three RNA bases mediate sequence-specific complex formation

To substantiate that the nucleotides U1780, C1779 and C1813 of *ASH1* E3 are important for base-specific LE recognition in solution, several mutants of E3 (51 nt) were created and their ability to form ternary complexes with She2p and She3p wild type was assessed in EMSAs (Figure 3.24). The mutation of U1780 to either the purine adenine in mutant "M1" or to another pyrimidine like cytosine in mutant "M2" impaired binding (Figure 3.24 A). This observation was substantiated by calculated corresponding apparent K_D values (Appendix Figure 5.14). While E3 (51 nt) wild type RNA formed the ternary complex wit a K_D of 0.2 ± 0.03 µM, the binding affinity of "M2" decreased 2-fold to a K_D of 0.44 ± 0.09 µM (Appendix Figure 5.14 A, E and Appendix Figure 5.14 D, E, respectively). Even worse was the apparent K_D of "M1" with a binding affinity of 1.05 ± 0.03 µM (Appendix Figure 5.14 C, E).
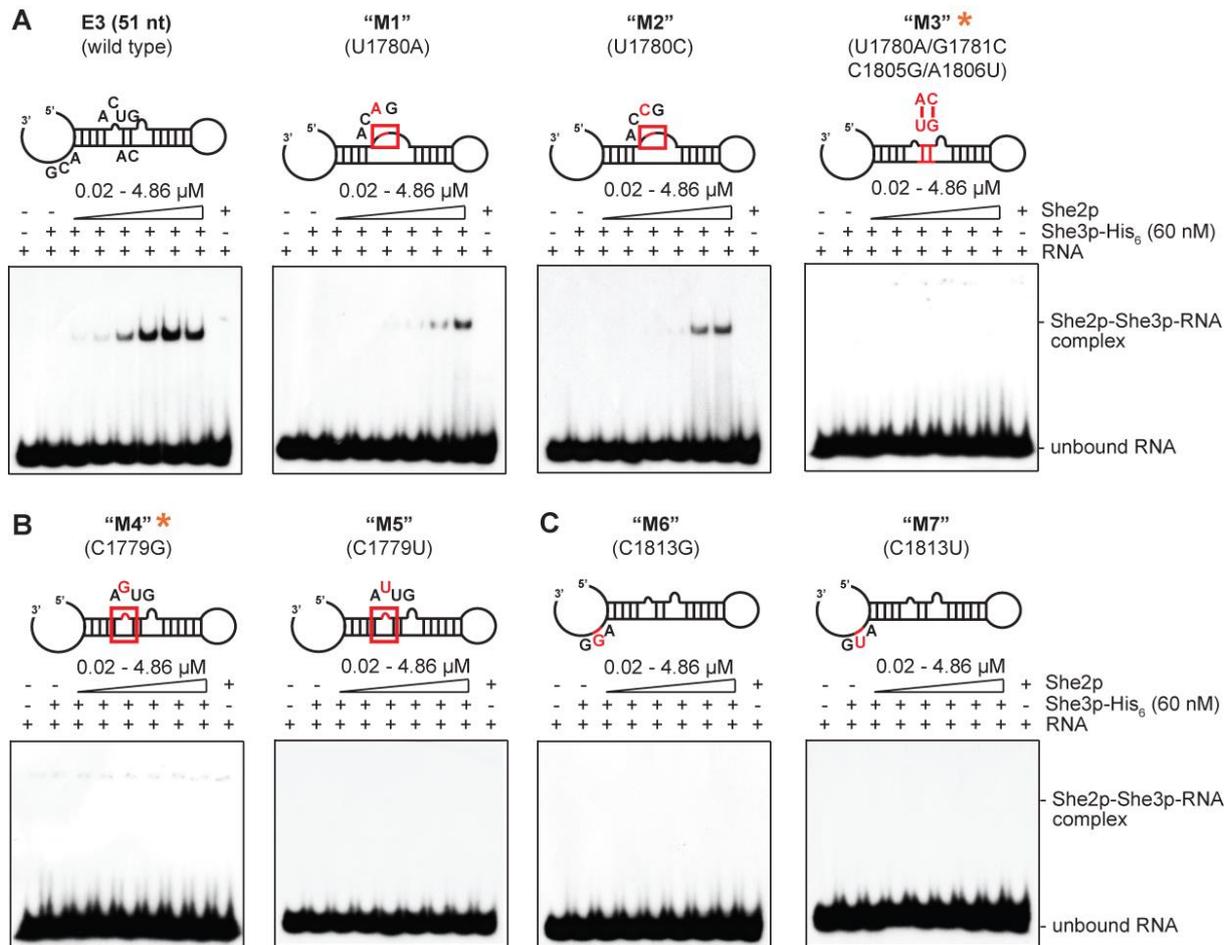
Figure 3.24: EMSAs with *ASH1* E3 (51 nt) mutants confirm base specific interactions. She2p wt concentrations were 20 nM, 60 nM, 180 nM, 540 nM, 1.62 µM and 4.86 µM, while full-length She3p was constantly present. Secondary structure predictions of E3 (51 nt) variants are depicted with their mutations highlighted in red. Orange asterisks mark EMSAs performed by Dr. Roland Heym. **A)** U1780 mutants "M1" and "M2" show reduced ternary complex forming abilities compared to E3 (51 nt) wild type RNA. The complementary mutation "M3" leads do abrogated RNA binding. **B)** Mutation of the conserved cytosine C1779 to either G in mutant "M4" or U in mutant "M5" results in the loss of ternary complex formation. **C)** Mutation of the second conserved cytosine at position 1813 to the purine G in mutant "M6" or to another pyrimidine like U in mutant "M7" also leads to abrogated RNA binding.

Taken together, although not being essential for ternary complex formation *in vitro*, the presence of a uracil at position 1780 contributes to binding. Interestingly Dr. Roland Heym has already introduced mutant "M3" during his studies, but the observed loss of She2p-She3p-RNA complex formation could not be explained (Figure 3.24 A). By inspecting the RNA in its bound form one can see that this complementary exchange directly affects the newly formed intramolecular base interactions upon She2p binding (Appendix Figure 5.10 and Figure 3.12 in section 3.3.4.4). Mutation of C1805 to G introduces repulsive forces with A1783. Mutation A1806U results in weaker base stacking between A1777 and A1778. Finally, the affinity-reducing effect of U1780A is added, which in combination

leads to disruption of ternary complex formation. Conserved cytosines C1779 and C1813 showed the tightest coordination in the dimeric as well as in the ternary co-complex. In line with these observations also the single-base mutations of either of the cytosines ("M4" and "M5" in case of C1779, and "M6" and " M7" in case of C1813) lead to a complete loss of RNA binding (Figure 3.24 B, C). Taken together these results show that uracil 1780 has a rather supportive role in RNA binding, while both specifically bound cytosines are essential for LE recognition. In addition the base composition in the central bulged region has to be precisely orchestrated, as it is important for the intramolecular stabilization of the RNA in its kinked form.

3.3.6.2 E172 and F176 in the protruding helix of She2p are necessary for ternary complex formation

The necessity of the protruding helix in She2p to fulfill the interaction with She3p and to perform synergistic RNA binding has already been shown [139]. Furthermore its deletion reduced RNA binding *in vitro* and lead to abrogated *ASH1*-mRNA localization *in vivo* [139]. In the co-structures amino acids E172 and F176 of the protruding helix make special contacts. They specifically interact with U1780 in the She2p-RNA co-complex, as well as with Y345 of She3p in the ternary complex. To assess the impact of these residues in solution a She2p (E172A, F176A) double mutant was created. CD assessment proved that secondary-structure properties were maintained upon protein mutation (Figure 3.25 A). EMSAs with this mutant revealed complete loss of RNA binding (Figure 3.25 B), underlining the importance of the before-mentioned specific contacts. In summary the positions in the protruding helix that are responsible for ternary complex formation could be experimentally confirmed to be amino acids E172 and F176.
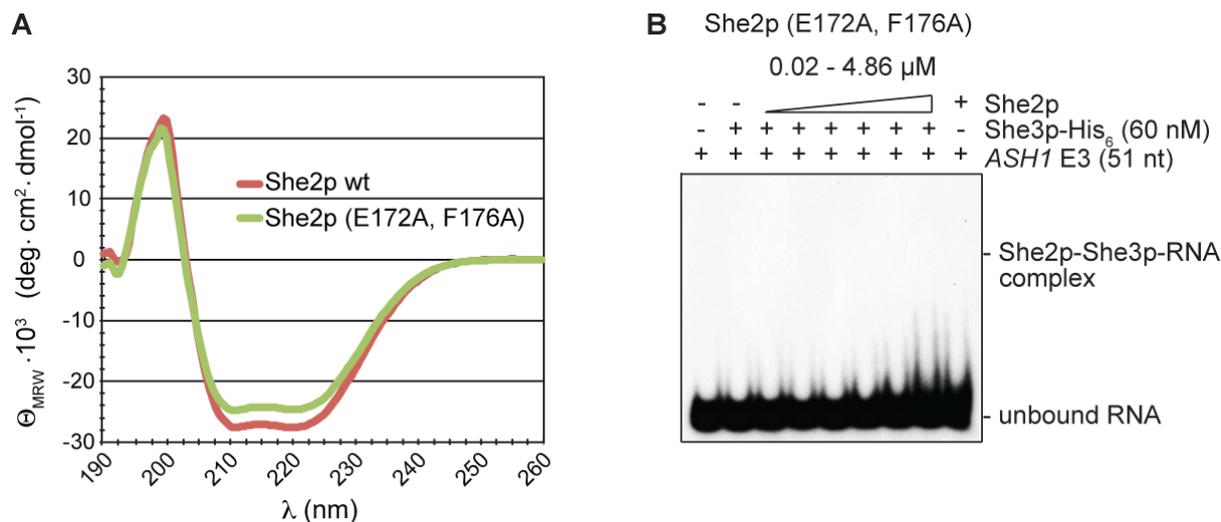
Figure 3.25: Analysis of mutant She2p (E172A, F176A) in solution. A) Far UV CD spectra showing that the alpha helical secondary structure of She2p is maintained upon mutation of E172A and F176A in the protruding helix. **B)** An EMSA with the double mutant She2p (E172A, F176A), She3p wild-type and E3 (51 nt) does not show any band shift arising from ternary complex formation. She2p concentrations were 20 nM, 60 nM, 180 nM, 540 nM, 1.62 µM and 4.86 µM.

3.3.6.3 Mutational analysis of the She3p R-site confirms binding mode in solution

The importance of the R-site of She3p in recognizing the RNA has been demonstrated by UV-crosslinking experiments [139] and the crystal structure of the ternary complex (section 3.3.5.10). Since the R-site intercalates between the protruding helix of She2p and the E3 LE, three important interaction sites were mutated and tested for ternary complex formation in solution (Figure 3.26). While the crystallized She2p-She3p fusion protein showed high affinity RNA binding (K_D = 112 ± 29 nM, Appendix Figure 5.15 C and Figure 3.26 A), the triple mutant (K340, R341 and Y345) showed a drastically reduced ability to bind RNA. A third EMSA where just the She3p R-site (residues 331-346) in the She2p fusion context was assessed for RNA binding highlighted that solely the R-site is sufficient for ternary complex formation, thereby emphasizing its importance for E3-LE recognition (Figure 3.26 A).

To evaluate the results in a more native background, EMSAs were also performed in a set-up that lacked the flexible linker (Figure 3.26 B, C). Here wild type She2p and She3p (331-405) were also able to complex the RNA although being less efficient. While in context of the She2p-She3p fusion-construct the RNA-binding ability of the She3p triple-mutant was reduced compared to wild type (Figure 3.26 A), EMSAs with the two individual proteins resulted in a complete loss of RNA binding (Figure 3.26 B, C).
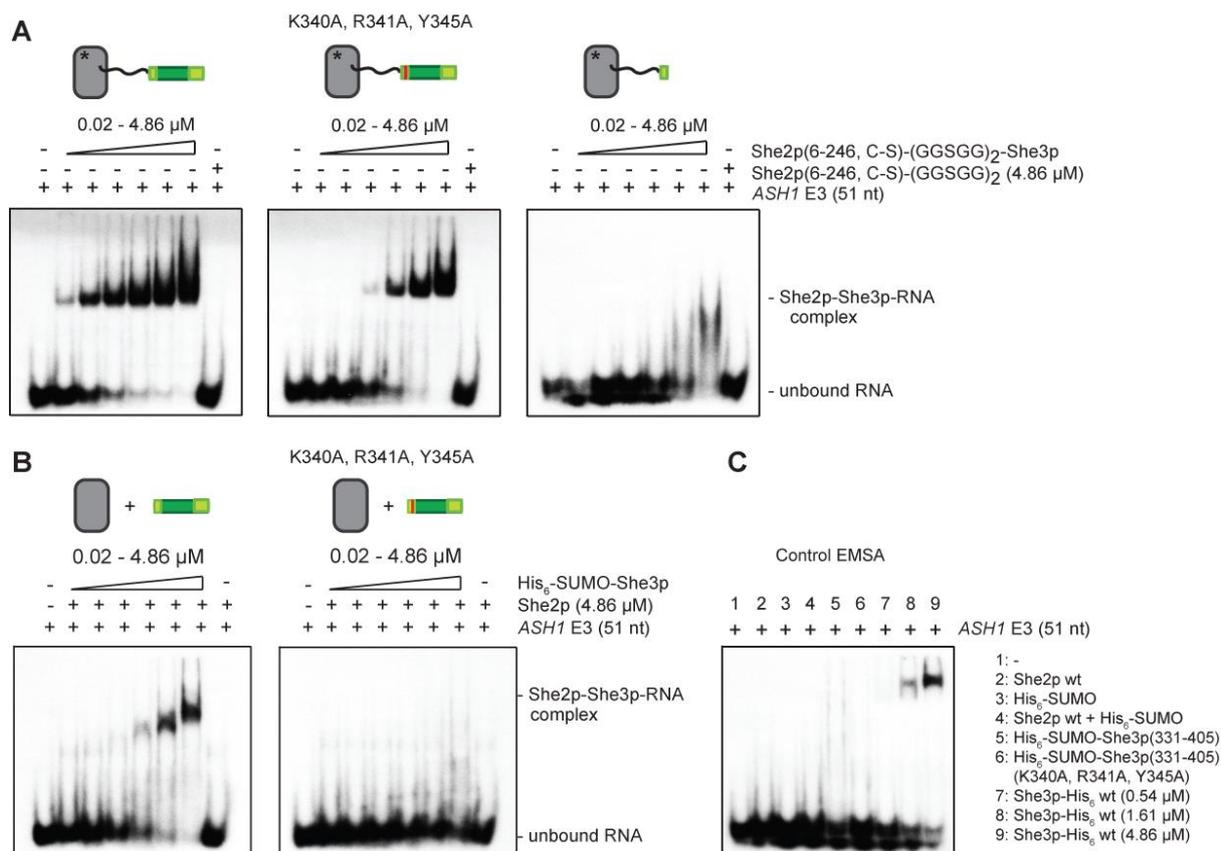
Figure 3.26: EMSAs with She3p variants confirms binding mode in solution. Schematic representations of She2p and She3p are generally depicted as in **Figure 3.19** A. Gray box without asterisk reflects She2p wild type. A red bar indicates She3p mutations. EMSAs were performed with radioactively labeled *ASH1* E3-(51 nt) RNA. Missing asterisks grey boxes in B) marks She2p wt. Increasing protein concentrations were 20 nM, 60 nM, 180 nM, 540 nM, 1.62 µM and 4.86 µM. **A)** The crystallized variant She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405) shows high affinity binding to E3 (51 nt) RNA (left gel), while the triple alanine mutant bearing key mutations K340A, K341A and Y345A has a drastically reduced RNA-binding ability (middle gel). The R-site fusion She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-346) is still able to form the ternary complex (right gel). **B)** EMSAs with wild type She2p and His$_6$-SUMO-tagged She3p (331-405) variants. While wild type She3p(331-405) forms a ternary complex with She2p and the RNA (left gel), this ability is lost upon mutation of K340A, K341A and Y345A (right gel). **C)** Control EMSA with proteins at a concentration of 4.86 µM unless stated otherwise. Single proteins or protein combinations used in B) are not sufficient to bind the E3-(51 nt) LE. Unspecific RNA binding is just observed for full-length She3p at very high concentrations.

To assess the binding interface of the She3p R-site more in detail, single amino acid mutantions of She3p (331-405) were created (K340A, R341A, R342A, F344A and Y345A) and their apparent binding with She2p for E3-(51 nt) RNA determined (Appendix Figure 5.15 A, B). Towards this aim, His$_6$-SUMO tagged She3p proteins were used since they had higher yields after purification. Distinct She2p-She3p-RNA band shifts could only be detected for wild type She3p (331-405) and mutant R342A. With an apparent K$_D$ of 3.43 ± 1.25 µM mutant R342A had a roughly 2-fold reduced affinity in ternary

complex formation compared to wild type She3p (K_D = 1.10 ± 0.32 µM). Since R342 is the only residue making intramolecular contacts, also this stabilizing interaction seems to be important for the RNA recognition of She3p. All other tested She3p single mutants (K340A, R341A, F344A and Y345A) did not show distinct band shifts and were therefore not quantified. Their binding affinities have to be worse than for R432A, underlining the importance of each of these single residues for ternary complex formation.

3.3.7 Specific contacts are essential for mRNA localization *in vivo*

In order to elucidate *in vitro* findings also *in vivo*, the impact of key interaction mutations in either She2p or She3p on *ASH1*-mRNA localization was assessed in living yeast. A system containing the GFP-tagged MS2-coat protein and an MS2-aptamer version of *ASH1* was used to track the RNA localization [134]. In multi-step cloning procedures I introduced several *SHE3* R- or/and P- site mutations to the yeast plasmid pRJ350 (FE#101-103 and FE#117-119, see Table 2.3). Additionally I performed initial cloning steps for the mutant She2p E172A, F176A (FE#104), which was finalized by Andreas Jenner and Muhammad Ibrahim Syed (Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany). The latter also conducted the *in vivo* experiments (Appendix Figure 5.16).

Yeast strains expressing She3p versions with different mutations in the R-site ((K340A, R341A) or (340-346A), respectively) or in the P-site ((L364A, V367A) or (364-368A), respectively) prevented proper mRNA localization (Appendix Figure 5.16 A). In line with these results also the double R- and P-site mutant She3p((340-346A)/(364-368A)) showed essentially abolished *ASH1* localization. Additional Western blot analysis was performed to test whether localization defects directly derive from the destroyed synergistic ternary complex upon She3p mutation or if the mutations had already an impact on protein expression (Appendix Figure 5.16 B). The majority of mutant yeast strains showed similar She3p expression levels as the wild-type strain. The sole exception was the triple mutant She3p K340A, R341A, Y345A, which was not expressed at all and was therefore also not quantified. *In vitro* the mutation of the R-site contact in She2p (E172A, F176A) destroyed synergistic ternary complex formation (Figure 3.25 in section 3.3.6.2). Also *in vivo* this mutant resulted in defective *ASH1*-mRNA localization. In summary both, the R- and the P-site of She3p are each essential for mRNA localization *in vivo*.

4 Discussion

4.1 Reconstitution of a functional mRNA-transport complex reveals mechanisms of assembly and motor activation

The assembly- and processivity-mediating features of motile particles for asymmetric mRNA localization are not well understood. Previous studies on mRNA-transport in yeast comprised experiments with RNP complexes purified from cell extracts [102], used partial complexes lacking the RNA cargo [164] or assembled complexes without knowing the exact stoichiometric requirements [159]. Recent *in vitro* reconstitution experiments clarified the stoichiometric ratios within a SHE complex [145] and provided a basis to study mRNP assembly and motility under well-defined conditions.

Since reconstitution experiments suggested equimolar concentrations of myosin and RNA in the SHE complexes [145], Dr. Dennis Zimmermann (University of Chicago, USA) was instructed to perform single-particle motility assays considering these requirements. Assembled particles containing RNAs with one *ASH1*-E3 LE indeed moved processively along actin *in vitro* and had a mean run length of 2.13 ± 0.89 µm (Appendix Figure 5.2 A, E, F). This is roughly twice as much as observed by Sladewski and colleagues who reported run lengths between 0.9 and 1.4 µm for single LE RNAs [159]. However, considering the distances that have to be traveled *in vivo* (~2-4 µm, depending on the size of the bud cell [199]), our result suggests that a significant fraction of SHE particles are likely able to move the whole distance from the mother cell to the bud tip of the daughter cell in one continuous run.

TIRFM experiments further showed that particles containing single LE RNAs had a mean velocity of 0.74 ± 0.20 µm/s (Appendix Figure 5.2 A, E, F). This observation is in close proximity to results from previous studies. For instance *in vitro* actin gliding assays revealed velocities for Myo4p ranging between 0.45 and 0.65 µm/s [160] and localizing particles *in vivo* were reported to move with velocities of 0.55 µm/s [199], 0.20-0.44 µm/s [134], and 0.63 µm/s [216].

The relevance of cargo RNA for complex assembly and its impact on particle motility has been controversially discussed [165], [159], [164]. To directly address this issue we compared particles with varying complex compositions. The tested experimental set-ups lacked RNA and either involved complexes that were composed of wild type components or comprised particles that carried an RNA binding-deficient She2p (N36S, R63K) mutant. Particle velocities of both experiments

(0.65 ± 0.15 µm/s and 0.62 ± 0.20 µm/s, respectively) as well as their run lengths (2.48 ± 1.02 µm and 2.02 ± 0.69 µm, respectively) were comparable with the results of the wild type set-up (velocity: 0.74 ± 0.20 µm/s, run length: 2.13 ± 0.89 µm) (Appendix Figure 5.2 E, F). Thus our findings pinpoint to an RNA-independent activation of particle motility *in vitro* and rather support a previous study from Krementsova and colleagues that suggested a negligible role for the RNA in motor activation [164]. Although the existence of RNA-free complexes *in vivo* is unlikely, these findings are valuable to deepen the mechanistic understanding of how a molecular motor is activated.

In vitro reconstitution experiments showed that *ASH1*-bound She2p associates with two She3p-Myo4p heterotrimers, which in turn assembles the SHE-core complex and induces Myo4p dimerization [145]. Since in our hands the cargo RNA is dispensable for processive movement, the necessity of the She2p-She3p interaction was challenged next. We found that particles reconstituted with the She3p-binding deficient mutant She2p ΔhE [139] were not able to move along actin filaments (Appendix Figure 5.2 F). This result confirmed that the protein-protein interaction between She2p and She3p is essential for proper SHE-complex assembly and its motility. Therefore I can conclude that it is rather the interplay between the RNA-binding proteins She2p and She3p that activates motility than the *ASH1*-mRNA cargo itself*.

ASH1 mRNA has not just one LE, but in total four elements [71], [61]. In a recent study by our lab, a combination of dynamic light-scattering analysis and sucrose gradient centrifugation was used to show that an RNA construct with two LEs results in large particles and indicated multimerization of SHE-core complexes [145]. To assess the potential effects of multiple motors on particle processivity, RNA with two LEs was subjected to single-particle motility assays. Photo-bleaching experiments showed twice as many fluorescence-intensity populations compared to the single LE-RNA containing particles, thereby confirming particle multimerization (Appendix Figure 5.3 D, E). The mean run length of these large particles (1.93 ± 0.68 µm, Appendix Figure 5.2 D-F) was essentially unchanged from the run length of particles containing single LE RNAs (2.13 ± 0.89 µm, Appendix Figure 5.2 A, E, F). The same was true for the mean velocities of both particles (single complex: 0.54 µm/s ± 0.13 µm/s versus oligomerized complex: 0.74 ± 0.20 µm/s). This finding again provides evidence against an RNA-based activation mechanism and supports the notion that neither more LEs, nor more motor molecules increase processivity.

However, our findings are in contrast to results from single particle motility assays performed by Sladewski and colleagues [159]. They showed that increasing the number of LEs in one RNA also

slightly extended the mean run lengths of SHE particles (single LE RNAs: 0.9-1.3 µm versus multi LE RNAs: 1.1-1.5 µm). These values were obtained by using a strong substoichiometric ratio of RNA over myosin (0.035 nM RNA and 25 nM myosin) and increased further by raising RNA and myosin concentrations (10 nM RNA and 250 nM myosin) to 1.4 µm for single LE RNAs and to 2.8 µm for a multi LE RNA [159]. This dramatic deviation from the correct stoichiometric ratios, which included an excess of motor molecules, could be one reason for the observed discrepancy. Nevertheless further technical differences could also be responsible for these contradicting results. While we used single F-actin filaments from rabbit skeletal muscles and wild type She2p for complex assembly, Sladewski and coworkers used stabilized yeast actin-fascin-tropomyosin bundles in combination with She2p as quadruple cysteine mutant. In summary, RNAs with multiple LEs and prolonged run lengths are unlikely to be essential for *in vivo* localization, but might act as modulators for efficient localization of RNAs.

Interestingly, also in *Drosophila* assemblies of high-order complexes with several *oskar* mRNAs have been found [217]. Here the dense packaging of *oskar*-RNP particles is thought to mediate translational repression. This observation is to some extent reminiscent of the large SHE particles that form *in vitro* [145] and could serve similar purposes in the active yeast mRNA transport.

Although in yeast we do not see any influence of cargo RNA on motor activation or processivity, a contrary effect is observed in the case of dynein-dependent transport in the *Drosophila* embryo. Here the dosage of LE-containing RNAs like *hairy* and the protein levels of BicD and Egl play an active role in regulating the initiation and maintenance of minus-end-directed mRNA transport [218], [219], [109].

Apart from *ASH1*-mRNA localization, She2p is involved in the transport of about 30 other transcripts into the daughter cell [111], [112], [113], [114]. In this respect the simultaneous transport of two different mRNAs (*ASH1* and *IST2*) is sometimes accomplished by the same particle *in vivo* [199]. We could recapitulate this observation in our reconstitution experiments *in vitro* and showed that particles including both Cy3.5-labeled *ASH1*-E3 and ATTO488-labeled *IST2* RNA moved actively along actin filaments (Appendix Figure 5.2 E, F). Since fluorophores were directly attached to the RNA, this experiment also served as direct proof that the reconstituted particles indeed transported RNA cargo.

In summary this study showed that run lengths of SHE complexes observed *in vitro* are compatible with the long-distance transport *in vivo*. The reconstituted particles offered similar motility properties, which were independent of the presence of RNA. This demonstrated that the RBP She2p and not its cargo activates motility. We further showed that SHE complexes have a defined size but multimerize into larger particles upon binding of RNAs with multiple LEs and are even able to bind different mRNAs.

4.2   Cargo RNA is dispensable for the formation of stable *ASH1* mRNP-core complexes

As described in the previous chapter, motility assays suggested that RNA is dispensable for complex assembly and the processive movement of SHE particles. This observation was contradictory to a previous study [159]. While Sladewski and colleagues claimed that the mRNA might be essential for motility, our results show the opposite. Moreover the authors stated in a recent commentary that particle movement in absence of RNA must have been artificially induced by protein-protein interactions at low ionic-strength (50 mM KCl) conditions and that "A stable Myo4p-She3p-She2p complex does not form at 140 mM KCl *in vitro* and thus Myo4p would be non-motile in the cell without cargo." [165]. Since no references were given to support this statement Annika Niedner-Boblenz and me directly assessed the stability of the *ASH1* mRNP-core complex in absence of cargo RNA at varying conditions.

Systematic pull-down experiments at different salt concentrations (50 mM, 140 mM and 200 mM sodium chloride) revealed stable and specific interactions of She2p-She3p and Myo4p in absence of RNA at 4 °C (Figure 3.4). Experiments at low ionic strength (50 mM sodium chloride) resembled conditions of single-particle motility assays and showed a significant portion of stably co-precipitated complex components. Since this ternary complex formation was absent when its She2p-She3p key interaction was destroyed by using the She3p-binding deficient mutant She2p (ΔhE) [139], we could show that the observed binding events indeed resemble specific interactions. These findings refute the above-mentioned statement of an artificial protein-protein interaction in absence of RNA [165] and strengthen the results of our single-particle motility assays.

Moreover we could show that such a stable and specific complex is also formed at a more physiologic salt concentration of 140 mM and the optimal yeast growth temperature of 30 °C. These robust

binding events suggest that the interaction of She2p-She3p and Myo4p is stable enough to mediate the activation of motility also in absence of RNA and invalidate the unsubstantiated annotation of Sladewski and colleagues [165].

Previous results demonstrated that SHE complexes, which included the *ASH1* RNA, also formed at 200 mM sodium chloride and are therefore even more salt resistant [145]. Our findings extend this view and indicate that the RNA is not even necessary to form such stable and specific complexes at elevated salt levels of 200 mM sodium chloride. Results from pull-down experiments showed stable co-precipitation of the complex components (Figure 3.4) and were further supported by analytical size-exclusion chromatography and SLS experiments (Figure 3.5 A and B, respectively). The latter showed that all components eluted in one RNA-free complex and a sub-fraction thereof was still fully assembled after more than two hours of size-exclusion chromatography. Its molecular mass was determined to be 420 kDa, which is in agreement with the calculated molecular weight of 434 kDa. Again this complex did not form with She2p (ΔhE) pinpointing at its specificity.

Interestingly, such insensitivity to high salt concentrations in absence of RNA was also observed previously for different SHE sub-complexes. For instance the interaction of She2p and She3p was stable at 200 mM sodium chloride in a pull-down and size-exclusion chromatography experiment [139] and Myo4p stably bound She3p in pull downs even at 1 M sodium chloride [163]. Together our data are consistent with results from previous work and underline the robustness and specificity of the She2p-She3p-Myo4p interaction in absence of RNA.

4.3 Structural characterization of the SHE-core complex reveals architecture and dynamics of *ASH1*-mRNA recognition

Dedicated RBPs selectively recognize transcripts for specific mRNA localization. However, a detailed mechanistic understanding on these events remains elusive. This is on one hand due to the fact that the highly specific recognition of mRNAs observed *in vivo* can be rarely recapitulated *in vitro* [220]. One exception is the highly stable and specific *ASH1* mRNA recognition by the RBPs She2p and She3p. On the other hand *cis*-acting elements usually show no recognizable conservation on sequence or secondary structure level although they are transported by the same particle. This complicates the deciphering of specificity-mediating features in mRNPs and indicates that high-resolution structures are essential to unravel distinct recognition modes.

In this thesis work, I presented the first crystal structure of a LE from budding yeast. A modified GAAA-TL/TLR module [200], being attached to the central *ASH1*-E3 stem was prerequisite for its crystallization and structure determination. Similar crystallization modules have been successfully used in the past and lead to the identification of different RNA structures that helped to obtain a more complete molecular view on their functional behavior. These included for instance a CUG repeat RNA helix, which reflects part of an aberrant CUG repeat expansion that induces myotonic dystrophy [202] or the core domain of the Hepatitis C virus internal ribosomal entry site [221].

The structure of *ASH1*-E3 comprised an elongated stem loop with three unpaired bases that pointed towards the solvent (C1779, A1782 and A1783, Figure 3.9 E, F). Interestingly Annika Niedner-Boblenz also highlighted these residues in previous chemical probing experiments [222]. Here the experimentally validated secondary structure suggested flexibility and likely structural "breathing" for bases C1779, A1782, A1783, and A1804/ C1805 on the opposing strand. These observations also nicely correlated with our solution NMR and SAXS data of the E3 element, which were performed in collaboration with the laboratory of Prof. Dr. Michael Sattler (Appendix Figure 5.9 and Appendix Figure 5.11, respectively). In NMR dynamics in the central double bulged region was observed, allowing the RNA to adopt different conformations. However, the elongated RNA conformation turned out to be the predominant species of the unbound and otherwise flexible RNA in solution.

Interestingly, internal dynamics of an LE RNA were also observed in the SOLE of the *oskar* mRNA of *D. melanogaster*. Here NMR analysis revealed fast internal motions of the medial stem-loop element [74]. Based on their data the authors reasoned that this apical helical stem is in conformational exchange with a second minor conformation that is melted. Although it is absolutely unclear if the observed dynamics are functionally important, the authors speculated that these dynamics could probably have a temperature-dependent regulation function or serve as option to structurally adapt the LE for the recognition of different protein partners during *oskar* transport [74].

Using the *ASH1* mRNP in yeast, I could answer for the first time the so far unresolved question how an RBP like She2p specifically recognizes one of its LEs. Towards this goal I first solved the crystal structure of the nuclear co-complex of She2p and *ASH1* E3 (Figure 3.10). One She2p tetramer binds two LEs on opposing sides, thereby formally establishing that the oligomeric state of the specifically RNA-recognizing She2p is indeed a tetramer. This is consistent with previous findings that assigned the tetrameric state to be required for RNA binding, mRNP assembly and mRNA localization [102], [152]. The structure also revealed that in contrast to the unbound RNA, LEs have to adopt a kinked

conformation upon She2p binding. This transition requires dramatic conformational changes (Figure 3.12), which could have not been predicted. Although an elongated LE reflects the major population of the unbound RNA, She2p recognizes the RNA in its kinked state. According to NMR and SAXS experiments this happens likely via conformational selection from a dynamic ensemble of different RNA-folding states, in which the central double bulged region acts as flexible hinge.

An intriguing feature in the co-structure is the point-symmetric She2p tetramer that binds asymmetric LE RNAs (Figure 3.10). This is possible since the RBP recognizes the RNA in a symmetric way. Each of the two opposing RNA-binding surfaces of She2p uses two identical sets of amino acids (N36, R52 and R63) to contact the conserved bases C1779 and C1813 in the asymmetric *ASH1* LE (Figure 3.11 A, B, C). Besides these two bases also U1780 specifically contacts two of the four protruding helices of She2p (Figure 3.11 A, D). My structural work shows that these three bases are the only specifically bound nucleotides in the nuclear She2p-RNA complex (Figure 3.11 and Figure 3.24). This observation nicely correlates with results from previous studies. For instance each single amino-acid mutant of She2p N36S, R52A and R63K completely abolished LE binding in filter-binding assays *in vitro* [150] and also mRNA localization *in vivo* [151], thereby underlining that these residues are indeed essential.

Interestingly slightly milder defects were observed for She2p mutants R43A, R44A and R49A *in vivo*. Depending on the mutant, *ASH1*-mRNA localization was reduced to levels between 4 and 50 % of wild-type localization [151]. Since the structure shows that these residues bind the RNA in a sequence-unspecific way (Figure 3.11), I conclude that they rather support initial LE recognition in the *ASH1* mRNA.

Apart from these residues that reside in the basic helical hairpin of She2p, also amino acids of its C-terminal tail contact the E3 (28 nt-loop) RNA. Amino acids H238, K243 and either G242 or L244 in E3 or E3', respectively, contribute to RNA binding at one side of the distal stem (Figure 3.11 A) This is consistent with previous experiments, in which a C-terminal deletion mutant (Δ241-246) showed reduced RNA binding *in vitro* and consistently loss of mRNA localization *in vivo* [139]. Although the last two residues were not resolved in the crystal structure, they could potentially help to recognize LEs with longer distal stem loops, as it would apply in case of E3 (51 nt) for instance.

In the co-structure the only contacts between the protruding helix of She2p and the RNA are E172 and F176 that interact with U1780 in E3 (Figure 3.11 A, D). Although the deletion of this protruding

helix resulted in a complete loss of ternary complex formation in EMSA experiments and abrogated the She2p-She3p interaction in pull downs, still a modest capability for RNA-binding remained in filter binging assays [139]. Thus, the protruding helix and U1780 seem to be rather required for the joining of She3p and only play a supportive role in the initial nuclear RNA recognition.

After nuclear export, the She2p-RNA co-complex is further extended by the Myo4p-bound She3p, which yields a highly stable, specific and motile mRNP. I succeeded in solving the SHE core complex structure of She2p, She3p, and the E3-RNA LE (Figure 3.22 A). It constitutes the very first high-resolution structure of a *cis*-acting LE in complex with its cognate protein complex of a mRNA-transport machinery. The She2p-She3p-RNA structure revealed that She3p indeed intercalates with its R-site between the protruding helix of She2p and the kinked RNA and thereby stabilizes their interaction. On each side of the She2p tetramer, two molecules of She3p symmetrically interact via their R-sites with two different regions of the RNA that are about 30 Å apart from each other (Figure 3.22 A). This dual RNA-binding mode is reminiscent of the symmetric LE recognition of She2p and provides additional sequence-independent RNA contacts that greatly decrease the spatial freedom of binding and result in strongly improved affinity and specificity in the ternary SHE core complex.

In contrast to the She3p R-site, which contacts both the RNA and She2p the P-site exclusively contacts She2p and does therefore not directly contribute to RNA recognition (Figure 3.23, Appendix Figure 5.12, Appendix Figure 5.13, Appendix Table 5.4). I also observed that the R-site buries a much bigger surface area on She2p than the P-site. Thus my findings indicate a rather supporting and complex-stabilizing role of the P-site. However, *in vivo* experiments clearly showed that both sites have to work together, since each of them was essential for mRNA localization (Appendix Figure 5.16). This fits the fact that both sites are highly conserved in She3p. The R-site includes a homology box region [340]KR(X)SFY(G/T)[346], and the P-site the conserved [364]LPG(V/L)(K/R)[369] hook motif [153]. The conservation indicates that a P-site supported, R-site mediated LE-recognition is potentially also present in other yeast species.

The ternary complex structure also explains previously reported UV cross-linking data. Here, an *ASH1*-E3 (51 nt)-fragment contacted She2p residues 164-179 [223]. Compared to the shorter E3-(28 nt-loop) RNA this fragment includes seven more bases at its 3' end. Considering the ternary complex structure with E3 (28 nt-loop), the anticipated continuation of its 3' end would be perfectly located to interact with the crosslinking region of She2p (Appendix Figure 5.17).

Comparing the cytoplasmic ternary complex with the nuclear She2p-RNA co-complex, the dual use approach of She2p residues N36, R52 and R63 in RNA recognition is maintained. In the ternary complex this concept is extended by She2p amino acids R43, K60 and She3p R341, making RNA binding even tighter. Also RNA flexibility, which is a prerequisite for the initial She2p binding, is necessary for the joining of She3p. U1780 that resides in the kink of E3 has to change its orientation towards the solvent to make space for the interaction between She3p Y345 and She2p F176 (Figure 3.22 D, E). Although She3p does not make base-specific contacts to the RNA, it stabilizes the phosphate backbone and introduces spatial restraints, which increase in stability and specificity in the ternary complex. In summary, the combined presence of two cytosines and one uridine in defined positions of an RNA-stem loop with mismatches in its stem are prerequisites for major RNA rearrangements and specific LE recognition by She2p and She3p.

These structural findings allow for the first time the deduction of features that precisely define a *cis*-acting LE. It is a combination of both structural and sequence constraints that were impossible to predict with bioinformatics tools. Such an interplay has not been observed so far and seems to be even more complicated as it has been proposed for other cases, such as the *Drosophila K10* TLS or the *oskar* SOLE stem loop. In both cases it is just known that proper mRNA localization does not require a distinct nucleotide sequence, but rather depends on the RNA secondary structure [73], [73].

For the SHE-protein dependent LEs in budding yeast, a previous study already identified the two cytosines C1779 and C1813 to be essential for mRNA localization and showed that they are also present in other LEs of the *ASH1* mRNA [224, 225] (Figure 4.1). By using computational search algorithms and secondary structure descriptors the researchers predicted a structural LE model in which the distance between the 3' phosphates of the two conserved cytosines was ~28 Å. However, their approach was not able to predict the actual binding mode that requires mismatches and structural flexibility of the RNA. My results show that the correct three-dimensional distance of the cytosine's 3' phosphates is about 38 Å when bound by She2p.

Interestingly another study found an adenosine six bases upstream of the short sequence motif they reported, also being present in a total of seven other LEs, including E1, E2A, E2B and E3 [143] (Figure 4.1). This adenosine corresponds to A1806 in the kinked region of E3 and is one of the bases that changes its interaction upon She2p binding (Figure 3.12). Although this base was not experimentally addressed by an isolated point mutation, a combinatorial mutant including A1806U resulted in

abolished ternary complex formation (Figure 3.24 A) and, according to the kinked RNA structure, in weaker predicted base stacking between A1777 and A1778 of the opposing strand. Even though these observations do not directly prove a certain sequence requirement in this particular position, they suggest that in addition to flexibility, also fine coordination of interactions could be important for the RNA to kink and thus also for RBP recognition.

In contrast to the conserved and essential cytosines that are found in all of the *ASH1* LEs, the uracil in position 1780 is lacking in other LEs (Figure 4.1). Since the mutation of U1780 in E3 decreased but did not abolish ternary complex formation in EMSAs (Figure 3.24 A), the substitution of its sequence position is likely tolerated in case of the other LEs. In summary there might be an alternative local binding mode in the other LEs in a position corresponding to U1780. This indicates that different subclasses of LEs might exist. Therefore further high-resolution structures will be required in the future to fully comprehend the features unique to all SHE-dependent LEs.
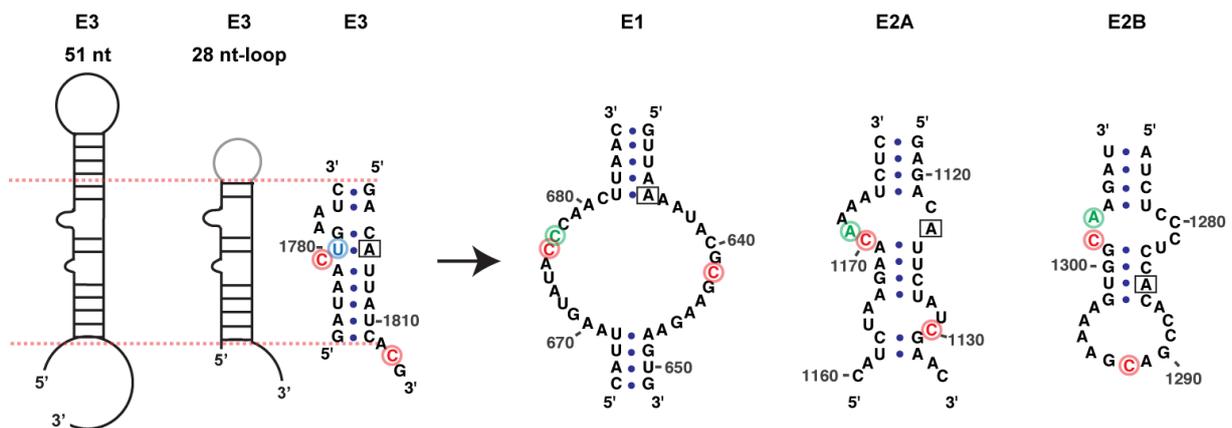
Figure 4.1 Comparison of predicted secondary structures from *ASH1* LEs. Essential bases in the *ASH1*-E3 LE that are specifically recognized by the She2p-She3p complex are highlighted in red (C1779/ C1813), and blue (U1780). Conserved cytosines are also found in E1, E2A and E2B. Based on secondary-structure predictions E2A and E2B are likely to have both cytosines in similar distance on the opposing strand of their stems. The base-specific position U1780 in E3 (highlighted in blue) is not conserved in the other elements (highlighted in green). Adenines marked with black boxes resemble an additional sequence pattern according to reference [143] and are found in all four LEs.

A new and interesting feature of this study is that conformational dynamics are a central feature in LE recognition and a prerequisite for the formation of the nuclear RNA-recognition and cytoplasmic transport complex. To date such RNA dynamics with functional importance have been mainly described for non-coding RNAs like ribosomal RNAs [226] and in particular for RNA elements with catalytic activity, like riboswitches [227] or the spliceosome [228]. To my knowledge, this is the first time that dynamics within a *cis*-acting LE are assigned to the actual function of a LE, i.e. the specific

recognition by RBPs. One unusual feature of the *ASH1* mRNA is that three out of four LEs reside in its open reading frame and only the E3 LE in the 3' UTR. One possible explanation for the dynamics in the LEs of *ASH1* is that rigid stem loops in the mRNA open reading frame would potentially act as roadblocks, hindering ribosomes from mRNA translation at the bud tip of the daughter cell. Therefore stem-loop flexibility could be a potential precondition to facilitate efficient translation at the site of transcript destination. Future work has to show if such dynamics in LE RNAs are indeed required for functions beyond their sole recognition by the transport machinery.

A recent proteome-wide study showed that a considerable number of RBPs contains disordered sequences. These intrinsically disordered regions (IDRs) allow for a large degree of flexibility and adaptability, which in turn facilitates the interaction with multiple binding partners [229]. However, increased RNA specificity was so far mainly attributed to the interplay of different (globular) RNA-binding domains that jointly recognize RNAs [220]. My results show that this principle can be extended to combinations of IDR-RBPs and globular RBPs. The unstructured C-terminal half of She3p seen in the crystal structure constitutes an IDR-RBP that binds the globular RBP She2p for specific RNA recognition. In case of She3p, the IDR-RBP modulates an existing binding event by the globular She2p to become more specific. It is tempting to speculate that several of the recently discovered IDR-RBPs [229] act in such combinatorial ways with globular RBPs to achieve higher affinity and specificity in RNA binding.

When the *ASH1* mRNP reaches the bud tip of the daughter cell, translational derepression causes locally restricted Ash1p translation. For this purpose the RBPs Puf6p and Khd1p are phosphorylated and thereby inactivated [169], [138]. [138]. Additionally it was reported that translational derepression is potentially actively supported by the phosphorylation of She3p [158]. Evidence derives from fluorescence *in situ* hybridization experiments and *SHE* genetic selection assays that showed that the phosphor-mimicking mutants She3p S348E and She3p (S343E, S361E) are defective in *ASH1*-mRNA localization and Ash1p sorting [158]. Interestingly, S343 and S348 reside in the She3p R-site, whereas S361 is located in its P-site. According to the ternary complex structure, S343 and S348 directly interact with several amino acids of She2p (Appendix Table 5.4, Appendix Figure 5.12). Furthermore, results from a previous study showed that She3p S348E reduces the ternary complex formation capability *in vitro* [139]. Together these data indicate that phosphorylation of the RBP She3p could indeed negatively regulate translational repression, possibly by destabilizing the mRNP and reducing its RNA-binding affinity. This is to some extent reminiscent of the modulation of the

RBP ZBP1, which is involved in the localization of *β-actin* mRNA in chicken embryo fibroblasts. Here the RNA-binding activity for ZBP1 is also negatively regulated by phosphorylation at the endpoint of mRNA transport [79].

In summary, this part of my study provided a comprehensive structural and functional analysis of how an mRNA-transport complex specifically recognizes its stem loop-containing LE RNA. I could structurally recapitulate the maturation process of an mRNP, beginning with the E3 LE in its unbound state via the nuclear She2p-RNA co-complex to the highly stable and specific cytoplasmic complex consisting of She2p, She3p and the RNA. I found that the *cis*-acting E3 LE of the *ASH1* mRNA alone adopts an elongated stem loop with a highly dynamic central bulged region between its flanking helices (Figure 4.2 – step 1). Upon co-transcriptional binding [146] She2p recognizes the RNA in its kinked state most likely via conformational selection from a dynamic ensemble of different RNA folding states. The transition from elongated to kinked state requires the RNA to undergo dramatic conformational changes (Figure 3.12). Therefore the first prerequisite for E3-LE recognition is RNA flexibility. Since the only specifically bound nucleotides in the binary complex are U1780 and the conserved cytosines C1779 and C1813, the second prerequisite is sequence identity in these positions (Figure 4.2 – step 2). In the nucleus this complex is stabilized by the RBP Loc1p, which is replaced at the nuclear pore by the cytoplasmic Myo4p-bound She3p by a mechanism that is not yet fully understood [168]. Additional joining of the unstructured She3p induces the formation of a synergistic cytoplasmic complex (Figure 4.2 – step 3, 4), in which a multi-interaction network of She2p and She3p recognizes a range of features in the E3 RNA (Figure 3.23). Since U1780 has to make space for the She3p R-site (Figure 3.22 D, E) only two sequence-specific bases (C1779, C1813) are recognized at defined positions in the ternary complex (Figure 3.23). My work demonstrated that the increase in specificity and affinity is mainly achieved by a more precise recognition of the RNA shape via the She3p R-site. For correct *in vivo ASH1*-mRNA localization however, also the P-site of She3p in its C-terminus is necessary. The observed combination of such structural and sequence requirements for RNA binding and the dramatic rearrangements of the RNA upon binding could have not been predicted before and therefore helped to advance our understanding of specific mRNA recognition for its transport.
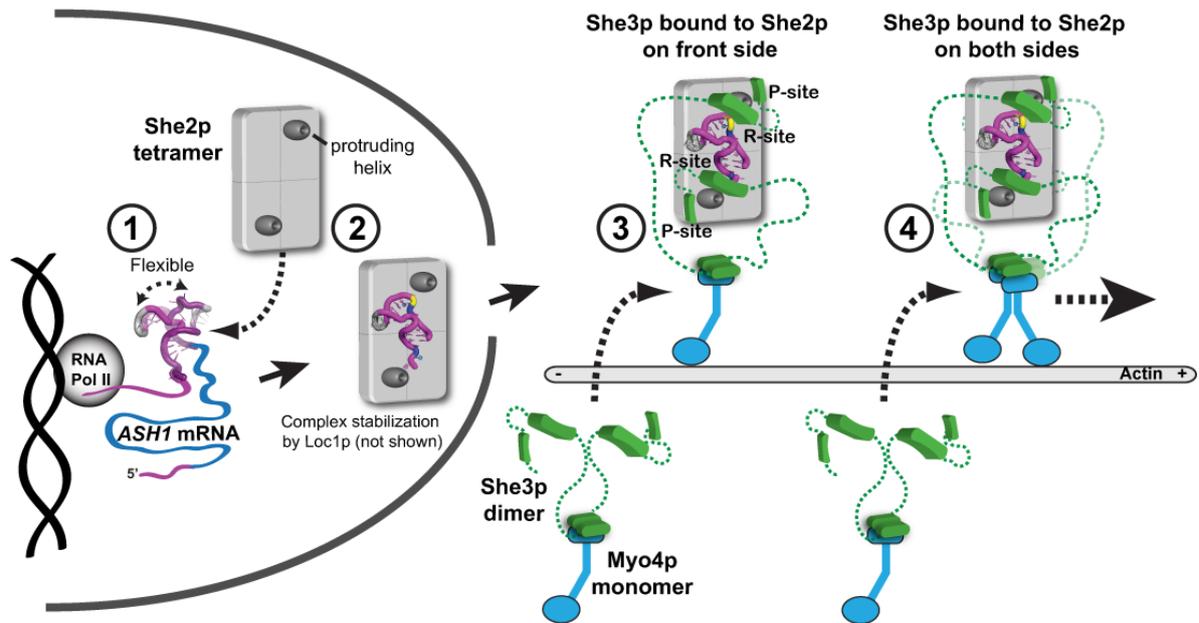
Figure 4.2: Summarizing model of specific *ASH1*-mRNA recognition and transport by She2p and She3p. Step 1: The E3 LE (pink) of the newly transcribed *ASH1* mRNA forms a stem loop with a dynamic bulge region in its stem that acts as flexible hinge. Although its major conformation is an elongated state different interconverting conformations are possible. **Step 2:** The tetrameric She2p (gray block) binds the E3 LE co-transcriptionally and E3 aligns in an L-like shape between the protruding helices (dark gray barrels) of She2p. This requires dramatic rearrangements in the bulged region of the RNA. The symmetric She2p specifically recognizes the asymmetric RNA via two conserved cytosines (dark blue) and an uracil (yellow). **Step 3:** After nuclear export the Myo4p-bound (blue) She3p (green) replaces Loc1p and joins the She2p-RNA complex. The association of She3p introduces mainly steric constraints that increase specificity and affinity for the LE RNA. The She3p R-site binds the RNA and She2p, while the P-site exclusively contacts She2p. **Step 4:** Since She3p binds on both sides of the She2p tetramer, two Myo4p molecules are incorporated into the complex, allowing for potential dimerization. As a result, active cargo transport can occur. Figure was modified from reference [222].

5   Appendix

	1	2	3	4
A	0.8 M Thiamine	40 % (w/v) Adonitol	15 % (w/v) Mannitol	10 mM $NaAuCl_4$
B	1.2 M Alanine	4 M γ-aminoburyric acid	1 M Ethanolamine-HCl	10 mM $HAuCl_4$
C	0.8 M Cysteine	40 % (w/v) Arabitol	1.5 M Glycyl-glycine-HCl	10 mM K_2PtCl_4
D	2 M Lysine	0.7 M Creatinine	30 mM Theophylline	10 mM $(NH_4)_2PtCl_4$
E	0.2 M Methionine	2 M Dimethylglycine	40 % (w/v) Fructose	10 mM $Pt(NH_3)_2(NO_2)_2$
F	2.5 M Serine	0.25 M Ammonium oxalate	1.5 M Na Glucuronate	10 mM K_2PtBr_4
G	1 M Arginine	0.5 M Itaconic acid	24 mM β-octyl glucoside	10 mM Magic triangle
H	1.5 M Acetylcarinitine	1 M Glucosamine-HCl	40 % (v/v) γ-butyrolaceton	1 mM Tantalum cluster

Table 5.1: Composition of the so-called heavy atom screen.

```
Conf: [||||||||||||||||||||||||||||||||||||||||]
Pred:  _____

Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
  AA: MSDQDNTQTSSSKLAPHHNIFMANLESSPTKDRNTSSQNA
              .         .         .         .
              10        20        30        40


Conf: [||||||||||||||||||||||||||||||||||||||||]
Pred:    ◖━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

Pred: CHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
  AA: SSSRVIESLHDQIDMLTKTNLQLTTQSQNLLSKLELAQSK
              .         .         .         .
              50        60        70        80


Conf: [||||||||||||||||||||||||||||||||||||||||]
Pred:   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

Pred: HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
  AA: ESKLLENLNLLKNENENLNSIFERKNKKLKELEKDYSELS
              .         .         .         .
              90        100       110       120


Conf: [||||||||||||||||||||||||||||||||||||||||]
Pred:   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

Pred: HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
  AA: NRYNEQKEKMDQLSKLAKNSSAIEQSCSEKLQNMEVNYNS
              .         .         .         .
              130       140       150       160


Conf: [||||||||||||||||||||||||||||||||||||||||]
Pred:   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━◗

Pred: HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCC
  AA: LLESQNLYRDHYSDEISKLNEKIGLLELELSNQNLNYGSD
              .         .         .         .
              170       180       190       200


Conf: [||||||||||||||||||||||||||||||||||||||||]
Pred:   ____◖━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

Pred: CCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
  AA: TSSNSDIELNLNKFNDSVKDLKSLETEKDSKLSKIITHSL
              .         .         .         .
              210       220       230       240


Conf: [||||||||||||||||||||||||||||||||||||||||]
Pred:   ◗─◖━━━━━━━━━━━━━━━━━━━━━━◗─◖━━━━━━━━━━━

Pred: HHCCHHHHHHHHHHHHHHHHHHHHHHHCCHHHHHHHHHHHH
  AA: DELNLQSWLNLYQTNENLISTFAEKMDLKDVLKRNDEKIS
              .         .         .         .
              250       260       270       280
```

Figure 5.1: Secondary structure prediction of She3p from *S. cerevisiae*. Region 331 to 405 is underlined in red. This part of She3p was successfully used in ternary complex crystallization and was predicted to be unstructured by PSIPRED v3.3 [186].

Figure 5.2: Single-particle motility assays with different SHE complexes visualized by TIRFM. A) Single-particle motility of SHE complexes in the presence of *ASH1*-3' RNA containing one E3 LE. **B)** Single-particle motility of the SHE complex in the absence of LE RNA. The presence of LE RNA does not greatly change the motility properties of the SHE complex. **C)** Single-particle motility assay of SHE complex with the RNA-binding deficient She2p (N36S, R63K) mutant confirms that the RNA cargo is dispensable for processive movement. **D)** Single-particle motility assays with RNA containing two LE elements does not show enhanced processivity. Mean run length and velocity of a tRNA-fusion construct with two *ASH1*-E3 LEs are in range of particles with only a single LE RNA. **E)** Box-plots representing data distributions for run lengths and velocities of each experiment, except for *ASH1*-3' and *IST2* RNA together. Here, the number of particles is too low for valid statistical analyses. Outliers (open circles, n < 28, for details see materials and methods in reference [145]) were determined to confirm the homogeneity in the collected data. **F)** Table summarizing results from the different TIRFM single-particle motility assays, in which the GFP from the GFP-She2p fusion protein was visualized (as indicated by the boxed molecule name in the complex descriptions), unless stated otherwise. The mean values for run length (left) and velocity (right) are derived from single-exponential and Gaussian fits to the distributions, respectively. n represents the number of processive runs analyzed in A)-E). Besides the motor and She proteins, particles were assembled in the presence of calmodulin, myosin-light chain, and the myosin-chaperone She4p. Experiments were performed and analyzed by Dr. Dennis Zimmermann. Movies and further experimental details can be inspected in [145].
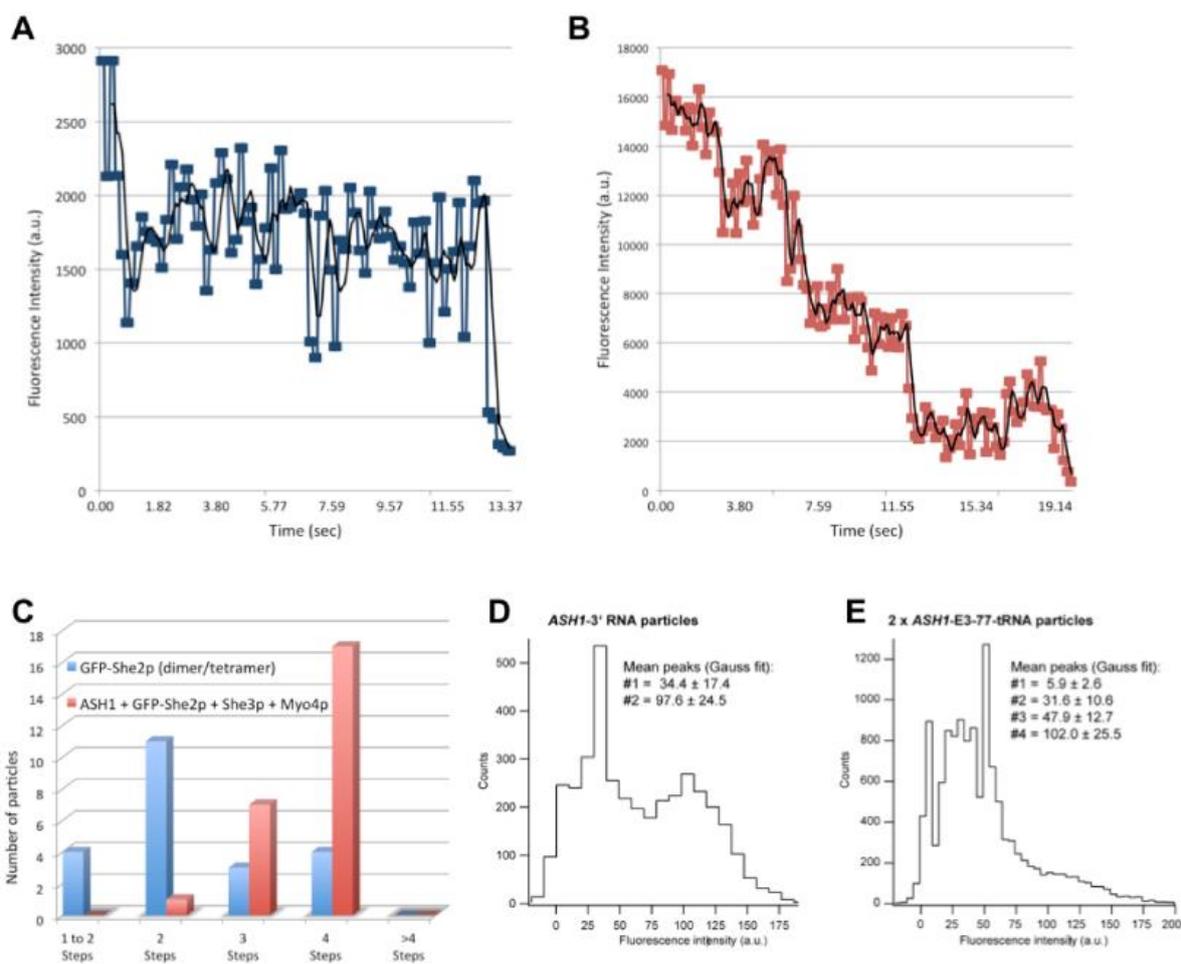
Figure 5.3: Photo-bleaching experiments. Shown are representative bleaching traces of **A)** GFP-She2p only and **B)** the reconstituted complex consisting of GFP-She2p, *ASH1*-3', full-length She3p and FLAG-Myo4p. **C)** Comparative plot of the number of bleaching steps per particle analyzed (n = 30). At the given experimental conditions (2 nM GFP-She2p (blue bars)), GFP-She2p exists in equilibrium between dimeric and tetrameric states. In the assembled motile complex (red bars), She2p becomes stabilized in its tetrameric state. **D)** and **E)** show histograms of fluorescence intensities from single-particle TIRFM experiments of *ASH1*-3' RNA particles and 2 x *ASH1*-E3-77-tRNA particles, respectively. Mean fluorescence intensity values are given for each population. While bleaching of *ASH1*-3' RNA particles yielded two populations (in D), 2 x *ASH1*-E3-77-tRNA particles revealed even four populations being present (E). Experiments were performed and analyzed by Dr. Dennis Zimmermann. Further experimental details can be inspected in [145].
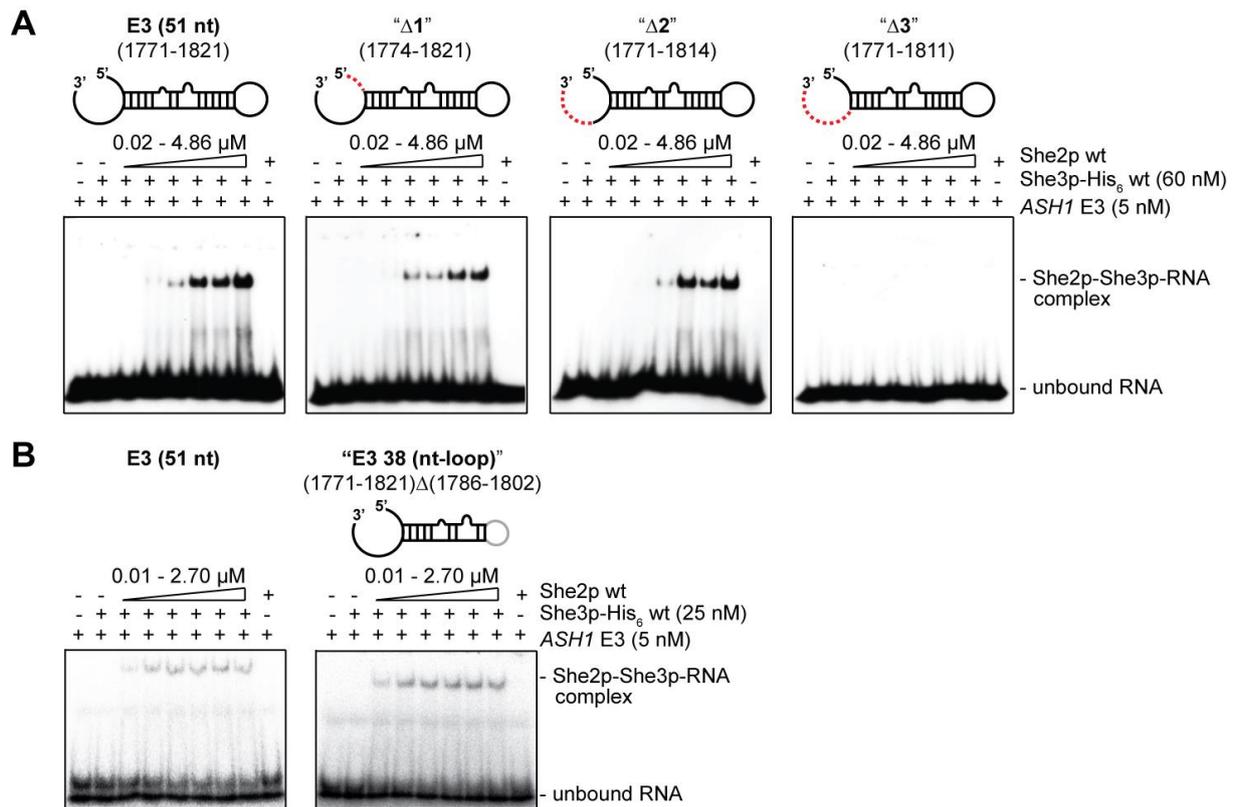
Figure 5.4: Assessment of synergistic RNA recognition by the She2p-She3p complex in electrophoretic mobility shift assays. A) Deletion of the single-stranded 5' and 3' regions in E3 (51 nt) (Δ1 and Δ2, respectively), resulted in wild-type binding. In contrast, a complete deletion of the single-stranded bases at the 3' end abolished binding (Δ3). E3 (51 nt) truncations Δ1 and Δ2 showed wild type-like complex formation. The results indicate that bases 1812-1814 are indispensable. Final She2p concentrations were 0.02 μM, 0.06 μM, 0.18 μM, 0.54 μM, 1.61 μM and 4.86 μM. **B)** A deletion of the upper part of the stem and of the nona-loop (nt 1786-1802) combined with an insertion of a more compact tetra-loop (38 nt-loop) still allowed for high-affinity binding by She2p and She3p. She2p concentrations were 0.01 μM, 0.03 μM, 0.10 μM, 0.30 μM, 0.90 μM and 2.70 μM. Results derive from Dr. Roland Heym.
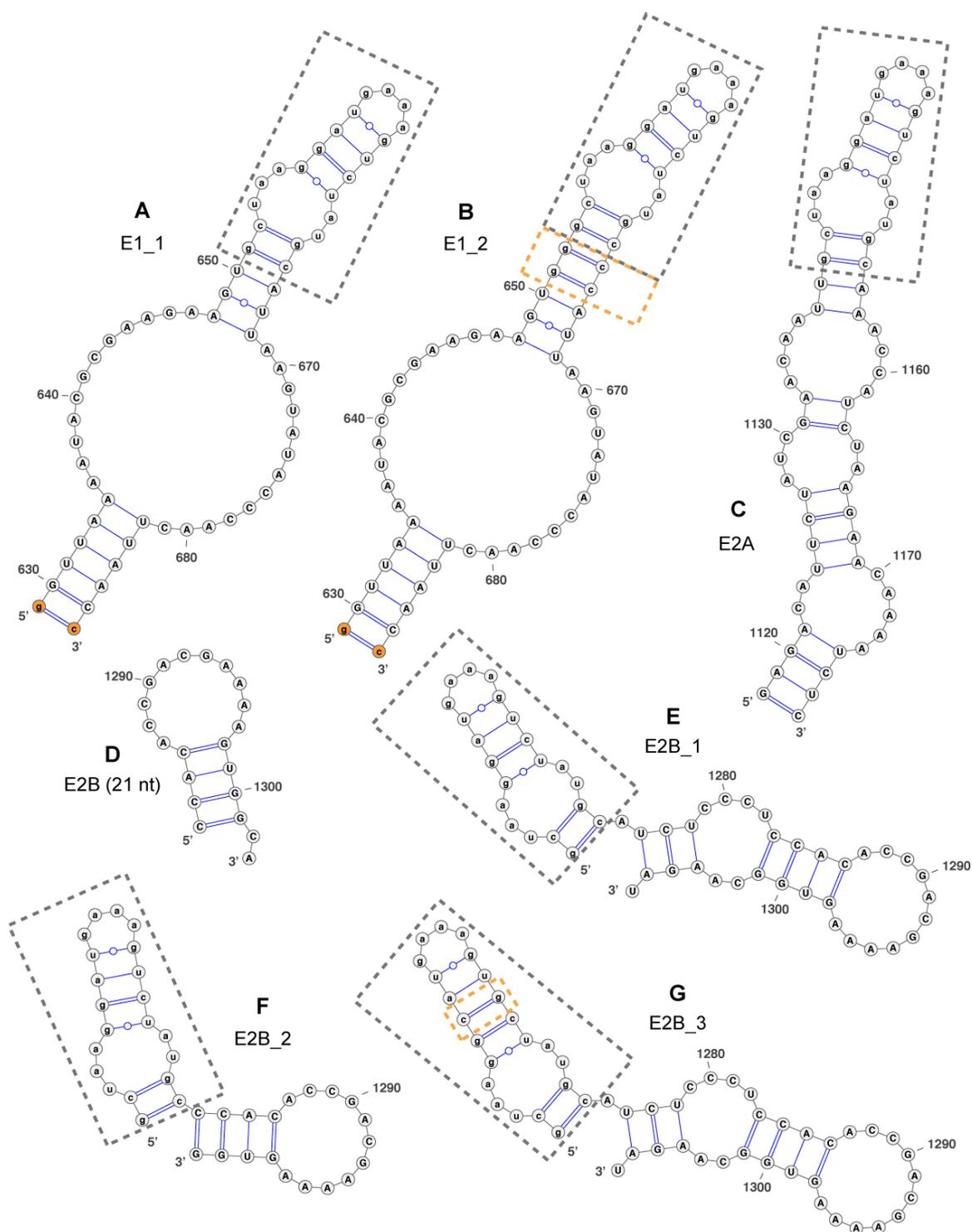
Figure 5.5: Secondary structure predictions of RNAs used for crystallization trials (1). Shown are the lowest initial free energy models of **A)** E1_1 (64 nt-TL/TLR), **B)** E1_2 (68 nt-TL/TLR), **C)** E2A (65 nt-TL/TLR), **D)** E2B (21 nt), **E)** E2B_1 (54 nt-TL/TLR), **F)** E2B_2 (40 nt-TL/TLR) and **G)** E2B_3 (56 nt-TL/TLR), predicted by mfold [144]. Grey, dashed box indicates bases forming the crystallization scaffold. Orange, dashed box marks base insertions in the scaffold. Small letters represent heterologous bases, naturally not belonging to the localization element. For clarification heterologous bases within a LE are accentuated in orange. Base numbering is related to the respective start codon.
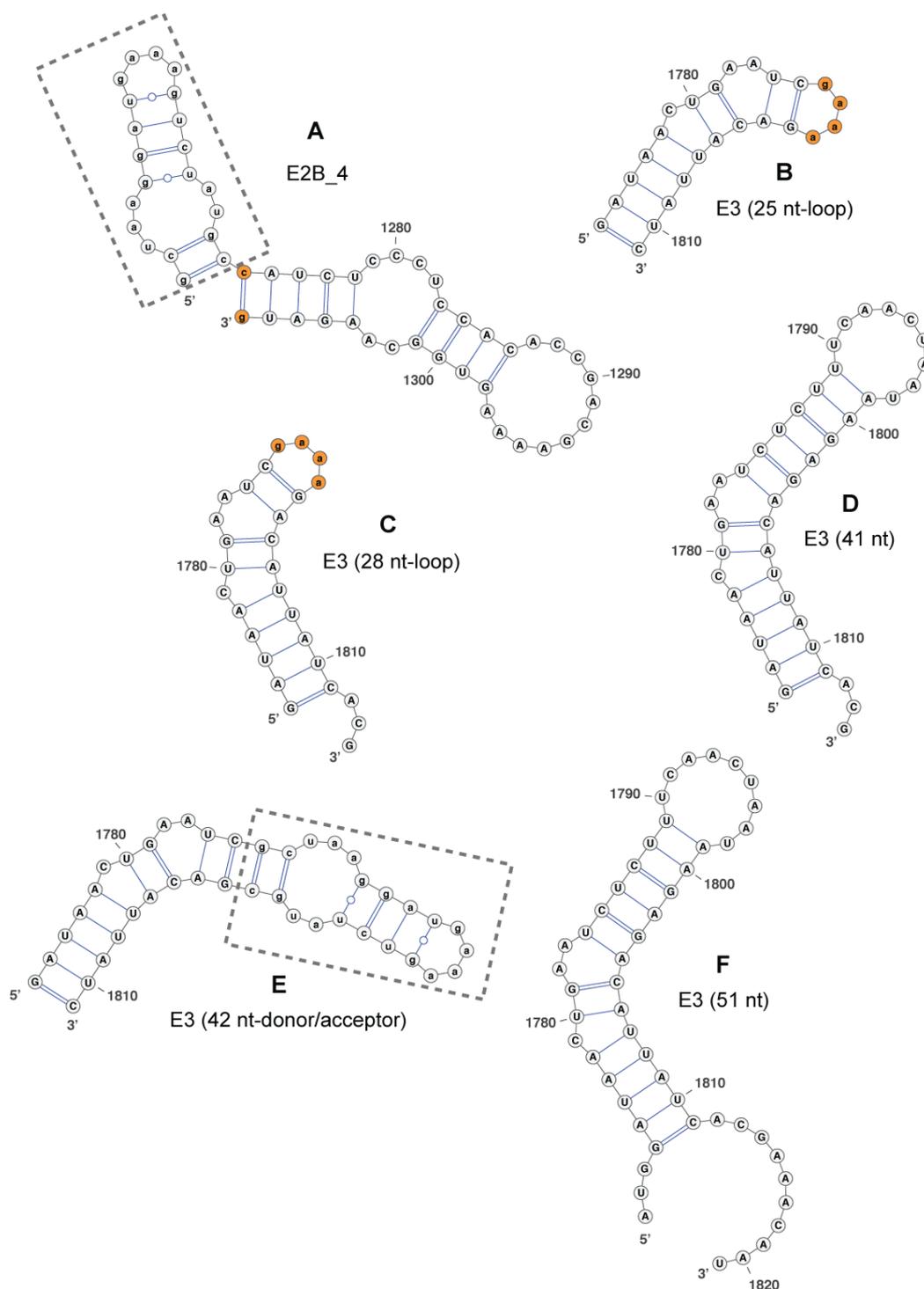
Figure 5.6: Secondary structure predictions of RNAs used for crystallization trials (2). Shown are the lowest initial free energy models of **A)** E2B_4 (56 nt-TL/TLR) **B)** E3 (25nt-loop) **C)** E3 (28 nt-loop) **D)** E3 (41 nt) **E)** E3 (42 nt-TL/TLR) **F)** E3 (51 nt), predicted by mfold [144]. Grey, dashed box indicates bases forming the crystallization scaffold. Orange, dashed box marks base insertions in the scaffold. Small letters represent heterologous bases, naturally not belonging to the localization element. For clarification heterologous bases within a LE are accentuated in orange. Base numbering is related to the respective start codon.
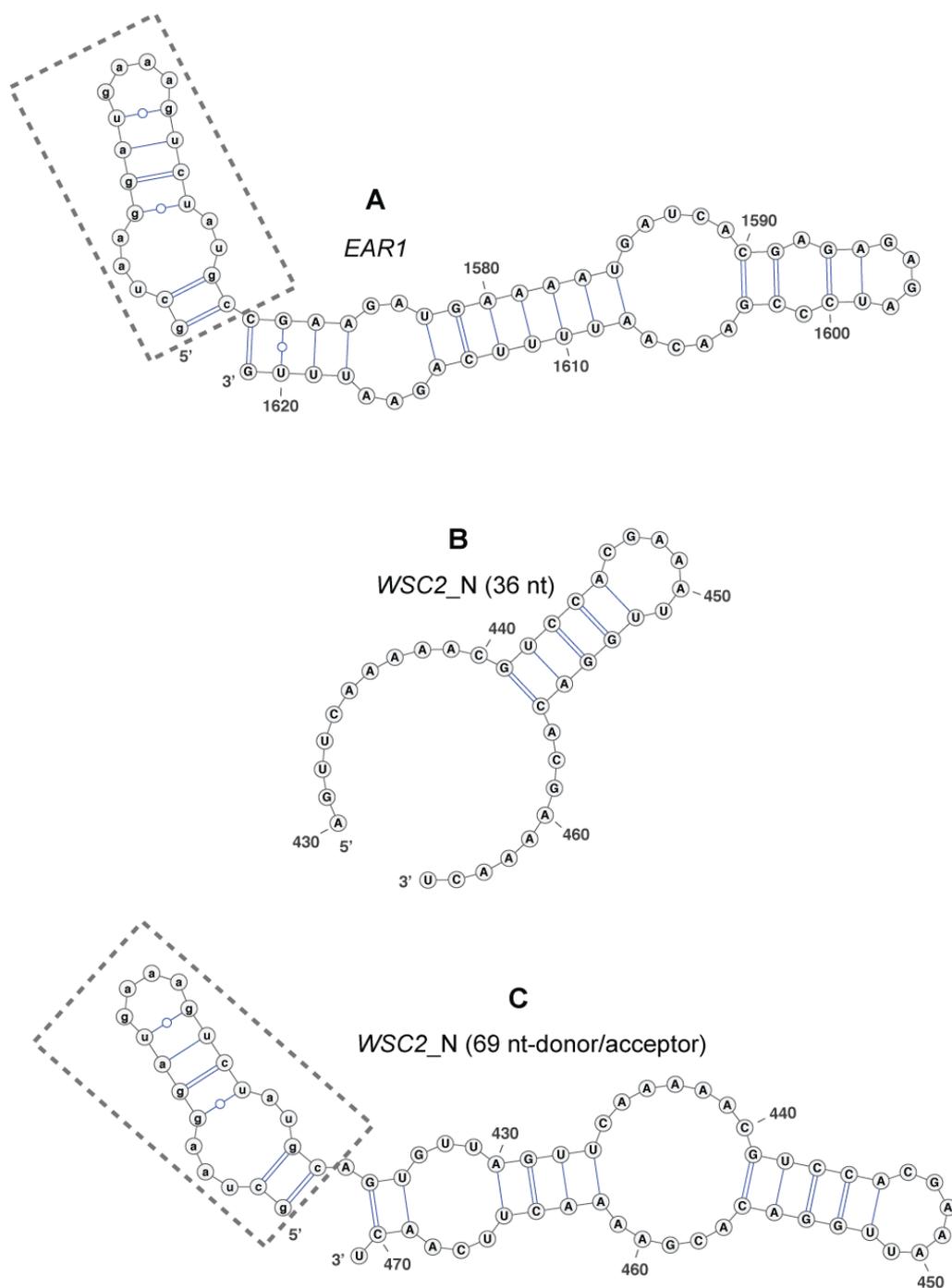
Figure 5.7: Secondary structure predictions of RNAs used for crystallization trials (3). Shown are the lowest initial free energy models of **A)** *EAR1* (71 nt-TL/TLR) **B)** *WSC2*_N (36 nt) **C)** *WSC2*_N (69 nt-TL/TLR), predicted by mfold [144]. Grey, dashed box indicates bases forming the crystallization scaffold. Small letters represent heterologous bases, naturally not belonging to the localization element. Base numbering is related to the respective start codon.

	RNA	RNA-She2p	RNA-She2p-She3p
Data collection/ processing			
PDB ID	5M0H	5M0I	5M0J
Beam line	ESRF ID23-2	ESRF ID 23-1	SLS X06DA (PXIII)
Wavelength (Å)	0.8726	1.0716	0.91745
Detector Distance (mm)	318.431	386.407	355.311
Number of images	47	612	3600
Oscillation range (°)	2	0.15	0.1
Space group	$P4_12_12$	$C2$	$C2$
Cell dimensions			
a, b, c (Å)	64.32, 64.32, 75.19	220.52, 58.34, 146.01	219.03, 58.98, 144.68
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 126.984, 90.0	90.0, 126.91, 90.0
Resolution (Å)	50-2.65 (2.72-2.65)*	100-2.41 (2.47-2.41)	100-2.80 (2.87-2.80)
R_{sym} or R_{merge}	7.3 (107.2)	6.8 (43.8)	18.9 (98.3)
$I / \sigma I$	19.93 (2.27)	10.37 (2.13)	10.34 (2.02)
Completeness (%)	99.6 (100.0)	90.1 (91.0)	99.6 (99.6)
Redundancy	7.24 (7.5)	1.9 (1.7)	6.8 (6.8)
Refinement			
Resolution (Å)	45.48-2.65	50.01-2.41	50.01-2.80
No. reflections	4,920	49,733	35,077
R_{work} / R_{free} (%)	21.1 / 24.0	18.9 / 23.5	19.4 / 24.6
No. atoms			
Protein	-	7,741	8,218
RNA	894	1,192	1,192
Ion	5	30	4
Ligand	-	44	-
Water	10	333	207
B-factors			
Protein	-	83.5[#]	60.1
RNA	91.9	59.2	46.8
Ions	49.0	64.8	59.4
Ligand	-	63.1	-
Water	67.7	50.5	29.7
R.m.s. deviations			
Bond lengths (Å)	0.01	0.02	0.01
Bond angles (°)	1.21	1.88	1.45
Ramachandran plot			
Allowed (%)	-	96.03	97.18
Additionally allowed (%)	-	2.14	2.19
Disallowed (%)	-	1.82	0.63

Table 5.2: Data collection/ processing and refinement statistics (molecular replacement) (1). *Values in parentheses are for highest-resolution shell. [#] Average B-factor for She2p was calculated to be 63.5, She3p had 110.2.

	She2p	She2p/She3p
Data collection/ processing		
Beam line	DESY P11 ID 29	ESRF ID 23-2
Wavelength (Å)	1.0332	0.8726
Detector Distance (mm)	612.724	317.623
Number of images	360	180
Oscillation range (°)	1	1.25
Space group	$P2_12_12$	$P2_1$
Cell dimensions		
a, b, c (Å)	115.46, 164.78, 65.83	102.82, 102.85, 142.65
α, β, γ (°)	90.0, 90.0, 90.0	90.00, 92.93, 90,00
Resolution (Å)	50-3.90 (4.00-3.90)	100-2.80 (2.87-2.80)
R_{sym} or R_{merge}	24.7 (94.0)	20.5 (81.8)
$I / \sigma I$	10.64 (3.15)	8.55 (2.49)
Completeness (%)	99.9 (100.0)	99.9 (99.9)
Redundancy	12.92 (0.01)	4.82 (4.85)
Refinement		
Resolution (Å)	47.32-3.90	142.46-2.80
No. reflections	11,421	69,587
R_{work} / R_{free} (%)	28.4 / 36.1	19.8 / 23.8
No. atoms		
Protein		15,212
Ion		8
Ligand		-
Water		608
B-factors		
Protein		52.2
Ions		41.7
Ligand		-
Water		24.3
R.m.s. deviations		
Bond lengths (Å)		0.01
Bond angles (°)		1.75
Ramachandran plot		
Allowed (%)		96.79
Additionally allowed (%)		2.71
Disallowed (%)		0.50

Table 5.3: Data collection/ processing and refinement statistics (molecular replacement) (2). *Values in parentheses are for highest-resolution shell. Refinement statistics for the She2p apo structure with a different space group are not complete since the refinement was not stretched to the limit.
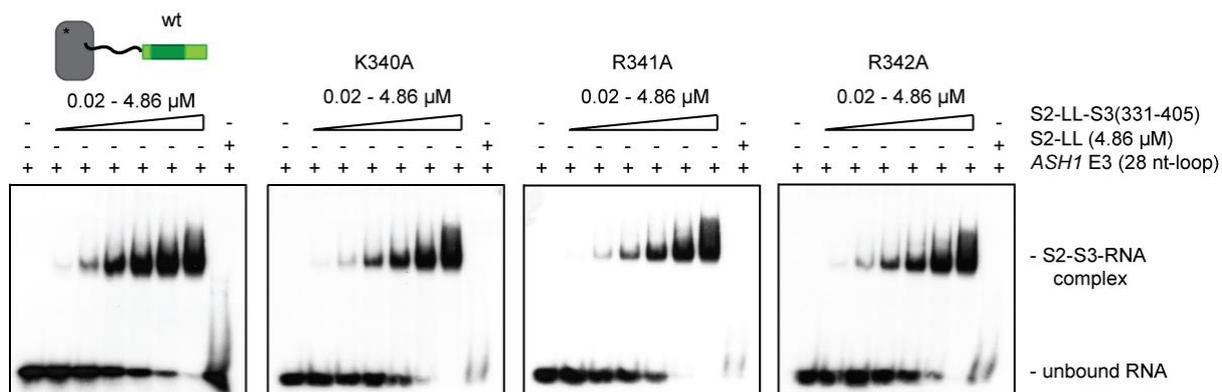
Figure 5.8: EMSAs of the E3 LE and selected alanine mutants in the longest She2p-She3p fusion construct background. Depicted are EMSAs with single amino acid mutations in She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405), termed "S2-LL-S3(331-405)" and radioactively labeled *ASH1*-E3 (28 nt-loop) RNA. Comparing K340A, R341A and R342A with the wild type construct shows reduced binding affinities for E3 RNA in each construct. Increasing protein concentrations were 20 nM, 60 nM, 180 nM, 540 nM, 1.62 µM and 4.86 µM. She2p(6-246, C-S)-(GGSGG)$_2$ termed "S2-LL" together with the minimal E3-LE served as control and did not show any band shift. EMSAs were performed in triplicates.
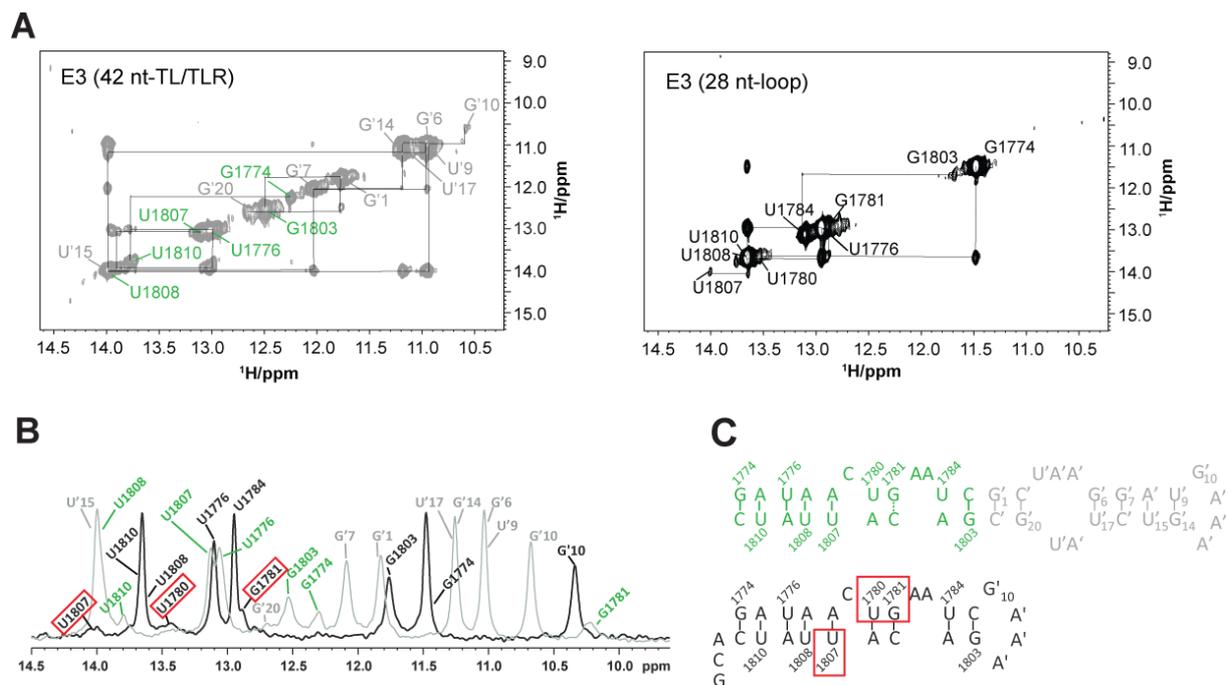
Figure 5.9: NMR analysis of E3 (42 nt-TL/TLR) and E3 (28 nt-loop) in solution. A) 2D-imino NOESY spectra of *ASH1*-E3 (42nt-TL/TLR) (left) and E3 (28 nt-loop) (right) RNAs showing the sequentially assigned observable imino protons. Resonance labels are color-coded according to the respective RNA fragments indicated in C). **B)** Overlay of 1D-imino proton-NMR spectra comparing the two RNA fragments. Line broadening of imino resonances and lacking signals in the bulged region indicate dynamics and flexibility in this part of the RNA. Bases of E3 (28 nt-loop) that show broadened peaks are highlighted with red boxes. **C)** Scheme of E3 (42 nt-TL/TLR) (top) and E3 (28 nt-loop) (bottom) RNA secondary structures based on NMR analyses. Red boxes reflect bases with line-broadened peaks from B). Solid lines in the upper scheme represent unambiguous H bonds as obtained from sequential assignments in A). Dotted line for G1781-C1805 in the 42-mer RNA indicates an assignment of the G1781 imino proton inferred by exclusion, although no cross peak to adjacent nucleotides was observed. Bases marked with an apostrophe indicate heterologous GAAA tetra-loop containing sequences, not belonging to the E3 LE. Experiments were performed and analyzed by Dr. Andreas Schlundt.
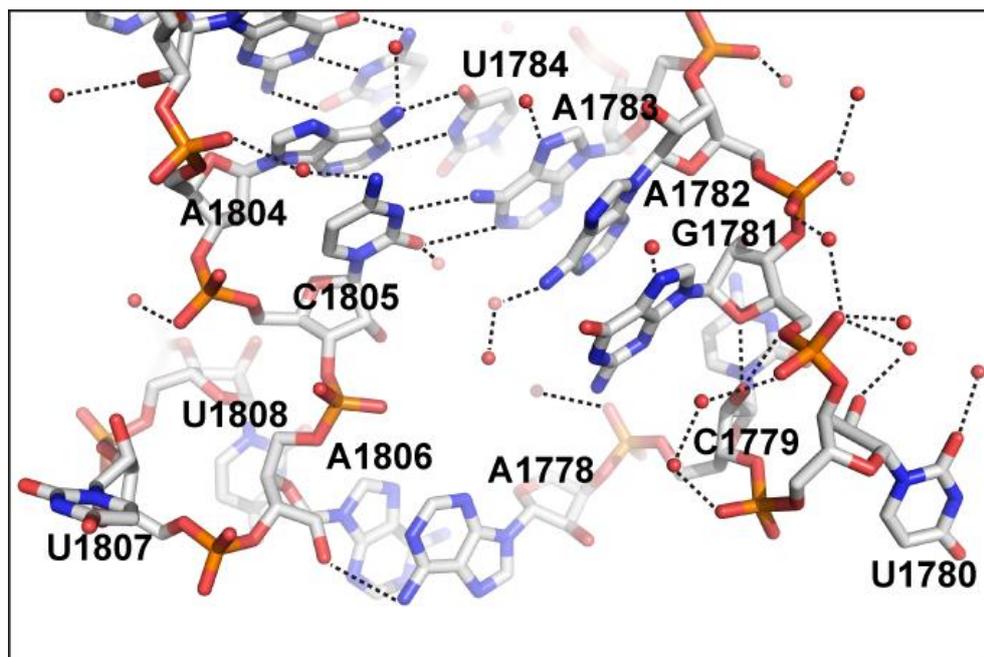
Figure 5.10: Close-up of interactions in the kinked region of E3 RNA in the She2p-bound state. Watson-Crick base pairings, hydrogen bonding, base stacking, and water interactions are shown. For better visualization, She2p is hidden.
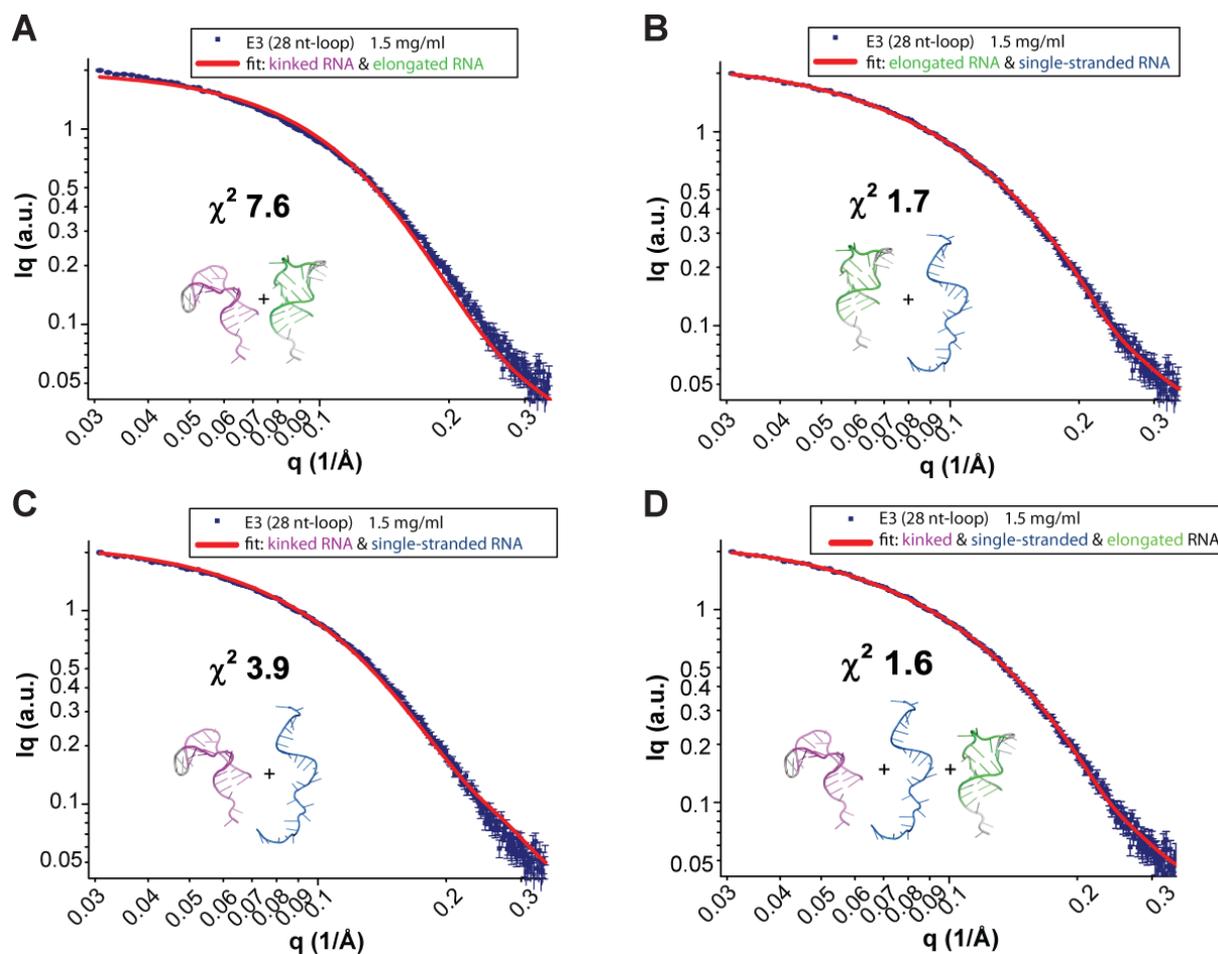
Figure 5.11: SAXS scattering curves of E3 (28 nt-loop) RNA. Back-calculated curves (red) for a mixture of different RNA models are fitted against the solvent-corrected scattering curve of E3 (28 nt-loop) RNA, recorded at a concentration of 1.5 mg/ml. Fit qualities are represented by χ^2 values. In **A)** a mixture of the bound (kinked, pink) and unbound (elongated, green) RNA species was tested. The calculated scattering curve of the fitted fractions consisting of kinked RNA : elongated RNA = 0.72 : 0.27 does not fit well to the recorded scattering curve, especially at low (0.03 Å^{-1}) and high (0.3 Å^{-1}) q values. This is also reflected by the high χ^2 value of 7.6. **B)** Assuming that an elongated and melted, single-stranded (ssRNA) species (depicted in blue; ratio elongated RNA : ssRNA = 0.65 : 0.34) exists in solution, the χ^2 value improves to 1.7. **C)** When testing a kinked and a single-stranded RNA model (ratio kinked RNA : ssRNA = 0.74 : 0.26) the fit is worse and χ^2 increases to 3.9. **D)** The most reliable results could be achieved by assessing a mixture of the kinked, single-stranded and elongated RNA. χ^2 improved to 1.6 (ratio kinked : ssRNA : elongated = 0.12 : 0.33 : 0.55).

She2p residue contacting the She3p R-site	Number of She2p chains with observed interaction	She3p R-site residue contacting She2p	Number of She3p chains with observed interaction
T53	4	K340	1
I56	4	R341	3
K57*	4	S343$^\$$	4
K60*	4	F344	4
D111	2	Y345	4
Y115	4	T346	4
T118	2	A347	4
Q119	4	S348$^\$$	4
I141	4	L351	2
L142	1		
D145	4		
D146	1		
E172	4		
F176	4		
E183	1		
N193	1		
I194	1		
F195[†]	1		
She2p residue contacting the She3p P-site	Number of She2p chains with observed interaction	She3p P-site residue contacting She2p	Number of She3p chains with observed interaction
W157	4	L364[†,§]	4
V173	1	P365[§]	4
N187	1	G366[§]	4
G189	1	V367[†,§]	4
E190	2	K368	3
T191	1	R369	2
D192	1	T370	2
I194	4	A371	2
L196[†]	2		
Q197[†]	4		
E198[†]	4		
I199[†]	4		
L200	4		
L211	3		
A214	3		
W215[§]	4		
I218	4		

Table 5.4: She2p-She3p R-/P-site protein-protein contacts. *Reference [150]: She2p mutation K57A decreased, and K60A abolished RNA-binding in filter-binding experiments. †Reference [139]: She2p (F195A, L196A), She2p (Q197A, E198A, I199A) and She3p (L364A, V367A) abrogated She2p-She3p interaction in pull-down experiments and reduced synergistic RNA binding in EMSAs. §Reference [153]: Single She3p mutations W215R, L364R, P365R, G366D and V367D showed severely reduced She2p-She3p binding in pull-down experiments. $Phospho-mimicking mutant She3p S348E showed moderately reduced synergistic RNA binding in EMSAs in reference [139] and mutants She3p S348E and She3p (S343E, S361E) delocalized *ASH1* mRNA *in vivo* in reference [158].
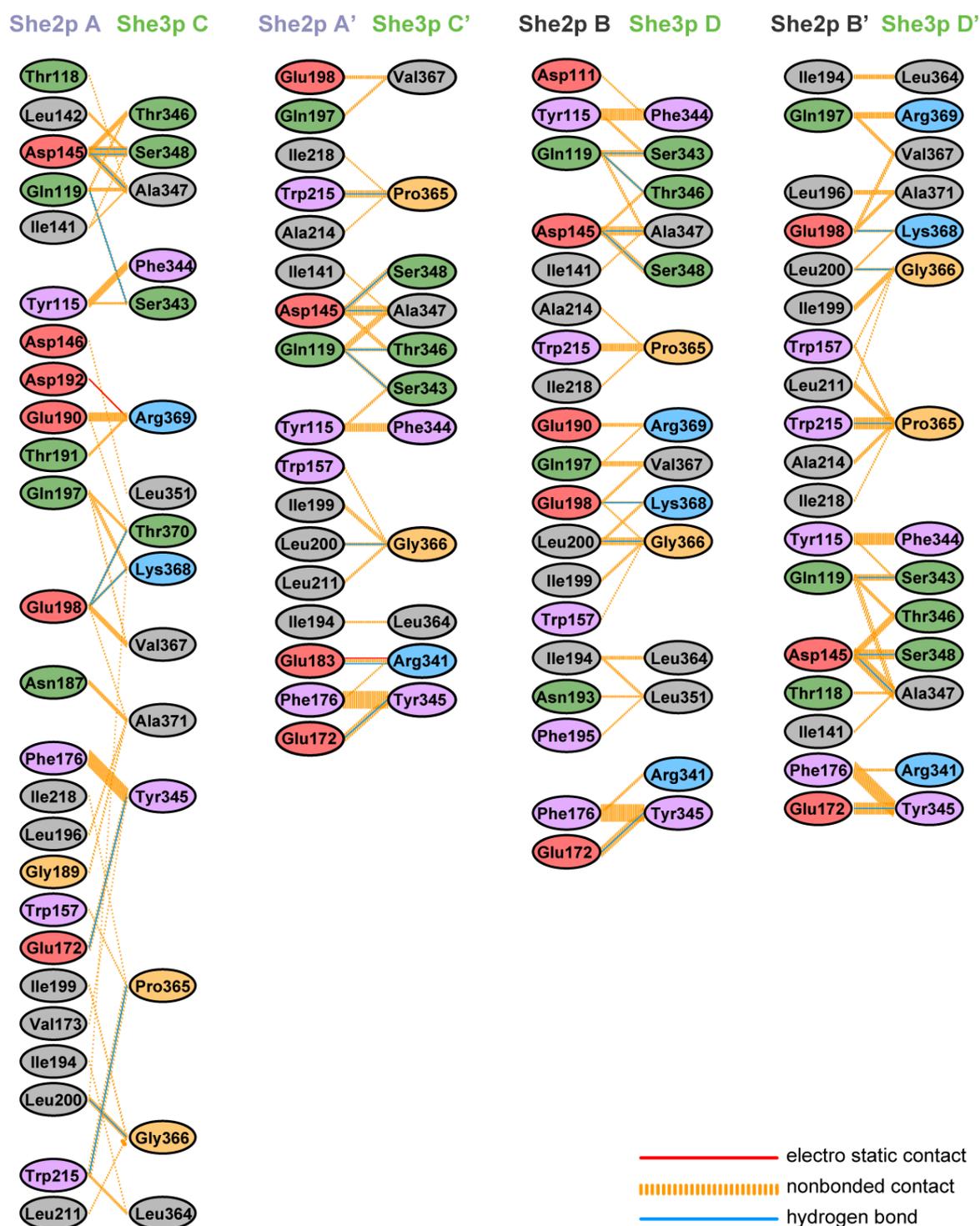
Figure 5.12: Protein-protein interactions of She2p and She3p in the ternary complex (1). The interactions between amino acids from different She2p chains (A, A', B and B') with their bound She3p chains (C, C', D and D') are indicated by colored lines. As shown at the bottom right side, those interactions are established by either electro static contacts (solid red lines), hydrogen bonds (solid blue lines) or non-bonded contacts (dashed yellow lines). The line broadness of the latter indicates the strength of interaction. Amino acids are shimmed according to the properties of their side chains: aliphatic (gray), aromatic (lavender), neutral (green), positive (red), negative (blue) and proline and glycine (orange). Interaction read-out was generated by PDBsum (section 2.12.12).

Figure 5.13: Protein-protein interactions of She2p and She3p in the ternary complex (2). The interactions between amino acids from different She2p chains (A, A', B and B') with their bound She3p chains (C, C', D and D') are indicated by colored lines. As shown at the bottom, those interactions are established by either hydrogen bonds (solid blue lines) or non-bonded contacts (dashed yellow lines). The line broadness of the latter indicates the strength of interaction. Amino acids are shimmed according to the properties of their side chains: aliphatic (gray), aromatic (lavender), neutral (green) and negative (blue). Interaction read-out was generated by PDBsum (section 2.12.12).

Figure 5.14 Representative EMSAs for apparent K_D determination of E3 LE-RNA variants. Binding affinities of E3 variants that showed ternary complex formation in **Figure 3.24** were quantified in EMSA experiments. Respective control EMSAs show unspecific RNA binding of She3p at high protein concentrations. Nevertheless, when She2p is additionally present, band intensities of the shifted complex increase, thus reflecting the specific ternary complex of She2p-She3p and RNA. Square brackets marked by an asterisk delimit the area of the gel, which was used for quantification. In **A)** radioactively labeled E3 (51 nt) RNA was used. In **B)** E3 (28 nt-loop) was assessed. **C)** and **D)** show EMSAs where RNA mutants "M1" and "M2" were tested. **E)** Table summarizing apparent mean K_D values ± s.d. for *ASH1* E3-RNA mutants in complex with She2p and She3p. K_Ds were calculated from three independent experiments using the one-site binding equation. While E3 (51 nt) forms the ternary complex with a K_D of 0.20 ± 0.03 µM, the minimal E3 (28 nt-loop) RNA bound She2p and She3p just slightly weaker. Replacing U1780 by the pyrimidine C in mutant "M2" decreased the affinity to a K_D of 0.44 ± 0.09 µM, whereas mutating U1780 to purine A in mutant "M1" had a more severe effect with a K_D of 1.05 ± 0.33 µM.

Figure 5.15: Representative EMSAs for apparent K_D determination of She3p variants. **A)** His$_6$-SUMO-She3p(331-405) and selected single-amino acid mutants were tested with She2p and *ASH1* E3 (51 nt) for their ternary complex-formation. Since distinct She2p-She3p-RNA complexes could be detected for wild-type She3p (331-405) and She3p (331-405) R342A apparent K_D values were determined. Square brackets marked by an asterisk delimit the area of the gel, which was used for quantification. **B)** Table summarizing calculated apparent mean K_D values ± s.d. for His$_6$-SUMO-She3p mutants. K_Ds were calculated from three independent experiments using the one-site binding equation. Binding affinities of She3p(331-405) mutated in positions K340A, R341A, F344A, and Y345A appeared to be worse than for R342A. These were not quantified due to the lack of distinct band shifts. **C)** Representative EMSA of the crystallized She2p(6-246, C-S)-(GGSGG)$_2$-She3p(331-405) fusion protein with *ASH1* E3-(28 nt-loop) RNA shows high affinity binding. The apparent K_D was 112 ± 29 nM. Square brackets marked by an asterisk delimit the area of the gel, which was used for quantification.

A

wild type

normal
ASH1- mRNA
localization

she3Δ

she3
R-site mutant

she3
(K340A, R341A)

defective *ASH1*-mRNA localization

she3
P-site mutant

she3
(L364A, V367A)

she3
R+P site

she2Δ

she2
(E172A, F176A)

defective *ASH1*-mRNA localization

	normal localization (%)	σ (%)
wild type	73.2	± 7.8
she3Δ	13.4	± 3.2
she3 R-site mut. (340-346A)	19.7	± 5.2
she3 (K340A, R341A)	25.5	± 7.7
she3 (K340A, R341A, Y345A)	n.d.*	-
she3 P-site mut. (364-368A)	12.9	± 3.3
she3 (L364A, V367A)	18.7	± 4.7
she3 R+P site mut.	15.7	± 3.4
she2Δ	23.0	± 0.6
she2 (E172A, F176A)	22.9	± 6.1

B

She3p-myc

Δ | wt | P-site mutant | R-site mutant | R+P double mutant

anti-myc ——— 55 kDa

anti-G6PD ——— 55 kDa

She3p-myc

Δ | wt | K340A, R341A | L364A, V367A | K340A, R341A, Y345A

anti-myc

anti-G6PD

She2p

Δ | wt | E172A, F176A

anti-She2p

anti-Pgk1

Figure 5.16: Assessment of specific *ASH1*-mRNA recognition by She2p and She3p *in vivo*. A) *In vivo* analysis of *ASH1*-mRNA localization using the MS2-GFP reporter system [134]. Arrowheads point at GFP- and therefore *ASH1*-mRNA containing particles. Different She2p and She3p mutant strains were compared to wild type. Exclusive localization of GFP-containing particles in the bud-tip was considered to be 'normal *ASH1*-mRNA localization', whereas exclusive localization in the mother cell or in both, the mother and the daughter cell, was considered to be 'defective'. Scale bar represents 2 μm. Pictures are overlays of DIC and GPF channels. **Table** (right) summarizes mean values in percent ± s.d. of budding yeast cells showing normal *ASH1*-mRNA localization of n independent cell culture replicates (3≤n≤9). Details and total cell numbers assessed per experiment can be inspected in Supplementary Data Set 1 of reference [222]. Asterisk indicates that for She3p K340A, R341A, Y345A no quantification (not determined (n.d.)) was done due to missing expression (see B). Experiments were performed by Andreas Jenner (Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany). **B)** Western blot comparing wt and She2p and She3p mutant expression levels in respective yeast strains used for *in vivo* experiments. Primary anti-myc antibody was used to detect C-terminally myc-tagged She3p. Anti-She2p antibody was used to detect She2p. Secondary anti-glucose-6-phosphat-dehydrogenase (G6PD) or phosphoglycerate kinase 1 (Pgk1) antibody was used to detect G6PD or Pgk1 levels, respectively, in the cell as measure for general protein expression. Apart from the expected She3p or She2p depleted strains, sole mutant She3p K340A, R341A, Y345A was not expressed at all. Experiments were performed by Muhammad Ibrahim Syed (Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany).
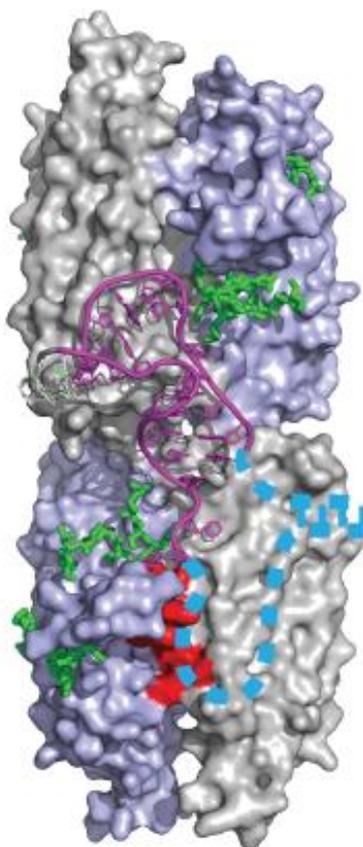
Figure 5.17: Ternary complex with She2p in surface representation. She2p monomers are depicted in light blue and gray. She3p R- and P-site residues are colored in green. E3 (28nt-loop) RNA is depicted in pink. Red area marks She2p amino acids 164-179 that were previously shown to UV crosslink with E3 (51 nt) [139]. Dashed blue line shows the anticipated projection of the RNA over the crosslinking site. This depicted single-stranded loop region followed by a double-stranded stem (right side) is consistent with the experimentally validated secondary structure from chemical probing experiments [222]

6 References

1. Martin, K.C. and A. Ephrussi, *mRNA localization: gene expression in the spatial dimension.* Cell, 2009. **136**(4): p. 719-30.

2. Medioni, C., K. Mowry, and F. Besse, *Principles and roles of mRNA localization in animal development.* Development, 2012. **139**(18): p. 3263-76.

3. Muench, D.G., C. Zhang, and M. Dahodwala, *Control of cytoplasmic translation in plants.* Wiley Interdiscip Rev RNA, 2012. **3**(2): p. 178-94.

4. Zarnack, K. and M. Feldbrugge, *Microtubule-dependent mRNA transport in fungi.* Eukaryot Cell, 2010. **9**(7): p. 982-90.

5. Nevo-Dinur, K., et al., *Translation-independent localization of mRNA in E. coli.* Science, 2011. **331**(6020): p. 1081-4.

6. Keiler, K.C., *RNA localization in bacteria.* Curr Opin Microbiol, 2011. **14**(2): p. 155-9.

7. Berleth, T., et al., *The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo.* EMBO J, 1988. **7**(6): p. 1749-56.

8. Lasko, P., *mRNA localization and translational control in Drosophila oogenesis.* Cold Spring Harb Perspect Biol, 2012. **4**(10).

9. Long, R.M., et al., *Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA.* Science, 1997. **277**(5324): p. 383-7.

10. Liao, G., et al., *Control of cell migration through mRNA localization and local translation.* Wiley Interdiscip Rev RNA, 2015. **6**(1): p. 1-15.

11. Misra, M., et al., *A Genome-Wide Screen for Dendritically Localized RNAs Identifies Genes Required for Dendrite Morphogenesis.* G3 (Bethesda), 2016. **6**(8): p. 2397-405.

12. Mikl, M., et al., *RNA localization in neurite morphogenesis and synaptic regulation: current evidence and novel approaches.* J Comp Physiol A Neuroethol Sens Neural Behav Physiol, 2010. **196**(5): p. 321-34.

13. Lecuyer, E., et al., *Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function.* Cell, 2007. **131**(1): p. 174-87.

14. Kim-Ha, J., J.L. Smith, and P.M. Macdonald, *oskar mRNA is localized to the posterior pole of the Drosophila oocyte.* Cell, 1991. **66**(1): p. 23-35.

15. Gavis, E.R. and R. Lehmann, *Localization of nanos RNA controls embryonic polarity.* Cell, 1992. **71**(2): p. 301-13.

16. Driever, W. and C. Nusslein-Volhard, *The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner.* Cell, 1988. **54**(1): p. 95-104.

17. Rebagliati, M.R., et al., *Identification and cloning of localized maternal RNAs from Xenopus eggs.* Cell, 1985. **42**(3): p. 769-77.

18. Birsoy, B., et al., *Vg 1 is an essential signaling molecule in Xenopus development.* Development, 2006. **133**(1): p. 15-20.

19. Takizawa, P.A., et al., *Actin-dependent localization of an RNA encoding a cell-fate determinant in yeast.* Nature, 1997. **389**(6646): p. 90-3.

20. Lawrence, J.B. and R.H. Singer, *Intracellular localization of messenger RNAs for cytoskeletal proteins.* Cell, 1986. **45**(3): p. 407-15.

21. Sundell, C.L. and R.H. Singer, *Actin mRNA localizes in the absence of protein synthesis.* J Cell Biol, 1990. **111**(6 Pt 1): p. 2397-403.

22. Condeelis, J. and R.H. Singer, *How and why does beta-actin mRNA target?* Biol Cell, 2005. **97**(1): p. 97-110.

23. Mili, S., K. Moissoglu, and I.G. Macara, *Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions.* Nature, 2008. **453**(7191): p. 115-9.

24. Trapp, B.D., et al., *Spatial segregation of mRNA encoding myelin-specific proteins.* Proc Natl Acad Sci U S A, 1987. **84**(21): p. 7773-7.

25. Muller, C., et al., *Making myelin basic protein -from mRNA transport to localized translation.* Front Cell Neurosci, 2013. **7**: p. 169.

26. Bassell, G.J., et al., *Sorting of beta-actin mRNA and protein to neurites and growth cones in culture.* J Neurosci, 1998. **18**(1): p. 251-65.

27. Holt, C.E. and S.L. Bullock, *Subcellular mRNA localization in animal cells and why it matters.* Science, 2009. **326**(5957): p. 1212-6.

28. Matrone, J., *How to move into the next era of health care.* Nurs Adm Q, 1996. **21**(1): p. 1-6.

29. Miller, S., et al., *Disruption of dendritic translation of CaMKIIalpha impairs stabilization of synaptic plasticity and memory consolidation.* Neuron, 2002. **36**(3): p. 507-19.

30. Besse, F. and A. Ephrussi, *Translational control of localized mRNAs: restricting protein synthesis in space and time.* Nat Rev Mol Cell Biol, 2008. **9**(12): p. 971-80.

31. Gu, W., et al., *Regulation of local expression of cell adhesion and motility-related mRNAs in breast cancer cells by IMP1/ZBP1.* J Cell Sci, 2012. **125**(Pt 1): p. 81-91.

32. Tolino, M., M. Kohrmann, and M.A. Kiebler, *RNA-binding proteins involved in RNA localization and their implications in neuronal diseases.* Eur J Neurosci, 2012. **35**(12): p. 1818-36.

33. Weis, B.L., E. Schleiff, and W. Zerges, *Protein targeting to subcellular organelles via MRNA localization.* Biochim Biophys Acta, 2013. **1833**(2): p. 260-73.

34. Karakozova, M., et al., *Arginylation of beta-actin regulates actin cytoskeleton and cell motility.* Science, 2006. **313**(5784): p. 192-6.

35. Mingle, L.A., et al., *Localization of all seven messenger RNAs for the actin-polymerization nucleator Arp2/3 complex in the protrusions of fibroblasts.* J Cell Sci, 2005. **118**(Pt 11): p. 2425-33.

36. Goley, E.D. and M.D. Welch, *The ARP2/3 complex: an actin nucleator comes of age.* Nat Rev Mol Cell Biol, 2006. **7**(10): p. 713-26.

37. Lin, A.C. and C.E. Holt, *Local translation and directional steering in axons.* EMBO J, 2007. **26**(16): p. 3729-36.

38. Simon, A.M., P. Hoppe, and S.J. Burden, *Spatial restriction of AChR gene expression to subsynaptic nuclei.* Development, 1992. **114**(3): p. 545-53.

39. Brenner, H.R., V. Witzemann, and B. Sakmann, *Imprinting of acetylcholine receptor messenger RNA accumulation in mammalian neuromuscular synapses.* Nature, 1990. **344**(6266): p. 544-7.

40. Forrest, K.M. and E.R. Gavis, *Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in Drosophila.* Curr Biol, 2003. **13**(14): p. 1159-68.

41. Zhou, Y. and M.L. King, *Localization of Xcat-2 RNA, a putative germ plasm component, to the mitochondrial cloud in Xenopus stage I oocytes.* Development, 1996. **122**(9): p. 2947-53.

42. Chang, P., et al., *Localization of RNAs to the mitochondrial cloud in Xenopus oocytes through entrapment and association with endoplasmic reticulum.* Mol Biol Cell, 2004. **15**(10): p. 4669-81.

43. Ding, D., et al., *Dynamic Hsp83 RNA localization during Drosophila oogenesis and embryogenesis.* Mol Cell Biol, 1993. **13**(6): p. 3773-81.

44. Bashirullah, A., et al., *Joint action of two RNA degradation pathways controls the timing of maternal transcript elimination at the midblastula transition in Drosophila melanogaster.* EMBO J, 1999. **18**(9): p. 2610-20.

45. Bashirullah, A., R.L. Cooperstock, and H.D. Lipshitz, *Spatial and temporal control of RNA stability.* Proc Natl Acad Sci U S A, 2001. **98**(13): p. 7025-8.

46. Semotok, J.L., et al., *Drosophila maternal Hsp83 mRNA destabilization is directed by multiple SMAUG recognition elements in the open reading frame.* Mol Cell Biol, 2008. **28**(22): p. 6757-72.

47. Semotok, J.L., et al., *Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early Drosophila embryo.* Curr Biol, 2005. **15**(4): p. 284-94.

48. Bonisch, C., et al., *Degradation of hsp70 and other mRNAs in Drosophila via the 5' 3' pathway and its regulation by heat shock.* J Biol Chem, 2007. **282**(30): p. 21818-28.

49. Dahanukar, A., J.A. Walker, and R.P. Wharton, *Smaug, a novel RNA-binding protein that operates a translational switch in Drosophila.* Mol Cell, 1999. **4**(2): p. 209-18.

50. Zaessinger, S., I. Busseau, and M. Simonelig, *Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4.* Development, 2006. **133**(22): p. 4573-83.

51. St Johnston, D., *Moving messages: the intracellular localization of mRNAs.* Nat Rev Mol Cell Biol, 2005. **6**(5): p. 363-75.

52. Gagnon, J.A. and K.L. Mowry, *Molecular motors: directing traffic during RNA localization.* Crit Rev Biochem Mol Biol, 2011. **46**(3): p. 229-39.

53. Jambhekar, A. and J.L. Derisi, *Cis-acting determinants of asymmetric, cytoplasmic RNA transport.* RNA, 2007. **13**(5): p. 625-42.

54. Giorgi, C. and M.J. Moore, *The nuclear nurture and cytoplasmic nature of localized mRNPs.* Semin Cell Dev Biol, 2007. **18**(2): p. 186-93.

55. Kress, T.L., Y.J. Yoon, and K.L. Mowry, *Nuclear RNP complex assembly initiates cytoplasmic RNA localization.* J Cell Biol, 2004. **165**(2): p. 203-11.

56. Kanai, Y., N. Dohmae, and N. Hirokawa, *Kinesin transports RNA: isolation and characterization of an RNA-transporting granule.* Neuron, 2004. **43**(4): p. 513-25.

57. Furic, L., M. Maher-Laporte, and L. DesGroseillers, *A genome-wide approach identifies distinct but overlapping subsets of cellular mRNAs associated with Staufen1- and Staufen2-containing ribonucleoprotein complexes.* RNA, 2008. **14**(2): p. 324-35.

58. Brown, V., et al., *Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome.* Cell, 2001. **107**(4): p. 477-87.

59. Heym, R.G. and D. Niessing, *Principles of mRNA transport in yeast.* Cell Mol Life Sci, 2012. **69**(11): p. 1843-53.

60. Kislauskis, E.H. and R.H. Singer, *Determinants of mRNA localization.* Curr Opin Cell Biol, 1992. **4**(6): p. 975-8.

61. Chartrand, P., et al., *Structural elements required for the localization of ASH1 mRNA and of a green fluorescent protein reporter particle in vivo.* Curr Biol, 1999. **9**(6): p. 333-6.

62. Cohen, R.S., S. Zhang, and G.L. Dollar, *The positional, structural, and sequence requirements of the Drosophila TLS RNA localization element.* RNA, 2005. **11**(7): p. 1017-29.

63. Pratt, C.A. and K.L. Mowry, *Taking a cellular road-trip: mRNA transport and anchoring.* Curr Opin Cell Biol, 2013. **25**(1): p. 99-106.

64. Munro, T.P., et al., *Mutational analysis of a heterogeneous nuclear ribonucleoprotein A2 response element for RNA trafficking.* J Biol Chem, 1999. **274**(48): p. 34389-95.

65. Hoek, K.S., et al., *hnRNP A2 selectively binds the cytoplasmic transport sequence of myelin basic protein mRNA.* Biochemistry, 1998. **37**(19): p. 7021-9.

66. Ainger, K., et al., *Transport and localization elements in myelin basic protein mRNA.* J Cell Biol, 1997. **138**(5): p. 1077-87.

67. Gao, Y., et al., *Multiplexed dendritic targeting of alpha calcium calmodulin-dependent protein kinase II, neurogranin, and activity-regulated cytoskeleton-associated protein RNAs by the A2 pathway.* Mol Biol Cell, 2008. **19**(5): p. 2311-27.

68. Chao, J.A., et al., *ZBP1 recognition of beta-actin zipcode induces RNA looping.* Genes Dev, 2010. **24**(2): p. 148-58.

69. Patel, V.L., et al., *Spatial arrangement of an RNA zipcode identifies mRNAs under post-transcriptional control.* Genes Dev, 2012. **26**(1): p. 43-53.

70. Ross, A.F., et al., *Characterization of a beta-actin mRNA zipcode-binding protein.* Mol Cell Biol, 1997. **17**(4): p. 2158-65.

71. Gonzalez, I., et al., *ASH1 mRNA localization in yeast involves multiple secondary structural elements and Ash1 protein translation.* Curr Biol, 1999. **9**(6): p. 337-40.

72. Subramanian, M., et al., *G-quadruplex RNA structure as a signal for neurite mRNA targeting.* EMBO Rep, 2011. **12**(7): p. 697-704.

73. Bullock, S.L., et al., *A'-form RNA helices are required for cytoplasmic mRNA transport in Drosophila.* Nat Struct Mol Biol, 2010. **17**(6): p. 703-9.

74. Simon, B., et al., *The structure of the SOLE element of oskar mRNA.* RNA, 2015. **21**(8): p. 1444-53.

75. Cheung, H.K., T.L. Serano, and R.S. Cohen, *Evidence for a highly selective RNA transport system and its role in establishing the dorsoventral axis of the Drosophila egg.* Development, 1992. **114**(3): p. 653-61.

76. Serano, T.L. and R.S. Cohen, *A small predicted stem-loop structure mediates oocyte localization of Drosophila K10 mRNA.* Development, 1995. **121**(11): p. 3809-18.

77. Ghosh, S., et al., *Control of RNP motility and localization by a splicing-dependent structure in oskar mRNA.* Nat Struct Mol Biol, 2012. **19**(4): p. 441-9.

78. Hachet, O. and A. Ephrussi, *Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization.* Nature, 2004. **428**(6986): p. 959-63.

79. Huttelmaier, S., et al., *Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1.* Nature, 2005. **438**(7067): p. 512-5.

80. Pan, F., et al., *ZBP2 facilitates binding of ZBP1 to beta-actin mRNA during transcription.* Mol Cell Biol, 2007. **27**(23): p. 8340-51.

81. Havin, L., et al., *RNA-binding protein conserved in both microtubule- and microfilament-based RNA localization.* Genes Dev, 1998. **12**(11): p. 1593-8.

82. Deshler, J.O., et al., *A highly conserved RNA-binding protein for cytoplasmic mRNA localization in vertebrates.* Curr Biol, 1998. **8**(9): p. 489-96.

83. Perycz, M., et al., *Zipcode binding protein 1 regulates the development of dendritic arbors in hippocampal neurons.* J Neurosci, 2011. **31**(14): p. 5271-85.

84. Song, T., et al., *Specific interaction of KIF11 with ZBP1 regulates the transport of beta-actin mRNA and cell motility.* J Cell Sci, 2015. **128**(5): p. 1001-10.

85. Gu, W., et al., *A predominantly nuclear protein affecting cytoplasmic localization of beta-actin mRNA in fibroblasts and neurons.* J Cell Biol, 2002. **156**(1): p. 41-51.

86. Lewis, R.A., et al., *Conserved and clustered RNA recognition sequences are a critical feature of signals directing RNA localization in Xenopus oocytes.* Mech Dev, 2004. **121**(1): p. 101-9.

87. Lewis, R.A., J.A. Gagnon, and K.L. Mowry, *PTB/hnRNP I is required for RNP remodeling during RNA localization in Xenopus oocytes.* Mol Cell Biol, 2008. **28**(2): p. 678-86.

88. Czaplinski, K., et al., *Identification of 40LoVe, a Xenopus hnRNP D family protein involved in localizing a TGF-beta-related mRNA during oogenesis.* Dev Cell, 2005. **8**(4): p. 505-15.

89. Czaplinski, K. and I.W. Mattaj, *40LoVe interacts with Vg1RBP/Vera and hnRNP I in binding the Vg1-localization element.* RNA, 2006. **12**(2): p. 213-22.

90.   Kroll, T.T., et al., *Interactions of 40LoVe within the ribonucleoprotein complex that forms on the localization element of Xenopus Vg1 mRNA.* Mech Dev, 2009. **126**(7): p. 523-38.

91.   Ramos, A., D. Hollingworth, and A. Pastore, *G-quartet-dependent recognition between the FMRP RGG box and RNA.* RNA, 2003. **9**(10): p. 1198-207.

92.   Bianco, A., et al., *Bicaudal-D regulates fragile X mental retardation protein levels, motility, and function during neuronal morphogenesis.* Curr Biol, 2010. **20**(16): p. 1487-92.

93.   Dienstbier, M., et al., *Egalitarian is a selective RNA-binding protein linking mRNA localization signals to the dynein motor.* Genes Dev, 2009. **23**(13): p. 1546-58.

94.   Hirokawa, N., et al., *Kinesin superfamily motor proteins and intracellular transport.* Nat Rev Mol Cell Biol, 2009. **10**(10): p. 682-96.

95.   Vale, R.D., *The molecular motor toolbox for intracellular transport.* Cell, 2003. **112**(4): p. 467-80.

96.   Ross, J.L., et al., *Processive bidirectional motion of dynein-dynactin complexes in vitro.* Nat Cell Biol, 2006. **8**(6): p. 562-70.

97.   Carson, J.H., et al., *Translocation of myelin basic protein mRNA in oligodendrocytes requires microtubules and kinesin.* Cell Motil Cytoskeleton, 1997. **38**(4): p. 318-28.

98.   Gagnon, J.A., et al., *Directional transport is mediated by a Dynein-dependent step in an RNA localization pathway.* PLoS Biol, 2013. **11**(4): p. e1001551.

99.   Jambor, H., et al., *A stem-loop structure directs oskar mRNA to microtubule minus ends.* RNA, 2014. **20**(4): p. 429-39.

100.   Mach, J.M. and R. Lehmann, *An Egalitarian-BicaudalD complex is essential for oocyte specification and axis determination in Drosophila.* Genes Dev, 1997. **11**(4): p. 423-35.

101.   Zimyanin, V.L., et al., *In vivo imaging of oskar mRNA transport reveals the mechanism of posterior localization.* Cell, 2008. **134**(5): p. 843-53.

102.   Chung, S. and P.A. Takizawa, *Multiple Myo4 motors enhance ASH1 mRNA transport in Saccharomyces cerevisiae.* J Cell Biol, 2010. **189**(4): p. 755-67.

103.   Zhang, H.L., et al., *Neurotrophin-induced transport of a beta-actin mRNP complex increases beta-actin levels and stimulates growth cone motility.* Neuron, 2001. **31**(2): p. 261-75.

104.   Nalavadi, V.C., et al., *Regulation of zipcode binding protein 1 transport dynamics in axons by myosin Va.* J Neurosci, 2012. **32**(43): p. 15133-41.

105.   Weingarten, M.D., et al., *A protein factor essential for microtubule assembly.* Proc Natl Acad Sci U S A, 1975. **72**(5): p. 1858-62.

106.   Drubin, D.G. and M.W. Kirschner, *Tau protein function in living cells.* J Cell Biol, 1986. **103**(6 Pt 2): p. 2739-46.

107.   Dixit, R., et al., *Differential regulation of dynein and kinesin motor proteins by tau.* Science, 2008. **319**(5866): p. 1086-9.

108.   McVicker, D.P., L.R. Chrin, and C.L. Berger, *The nucleotide-binding state of microtubules modulates kinesin processivity and the ability of Tau to inhibit kinesin-mediated transport.* J Biol Chem, 2011. **286**(50): p. 42873-80.

109.   Soundararajan, H.C. and S.L. Bullock, *The influence of dynein processivity control, MAPs, and microtubule ends on directional movement of a localising mRNA.* Elife, 2014. **3**: p. e01596.

110.   Singer-Kruger, B. and R.P. Jansen, *Here, there, everywhere. mRNA localization in budding yeast.* RNA Biol, 2014. **11**(8): p. 1031-9.

111.   Takizawa, P.A. and R.D. Vale, *The myosin motor, Myo4p, binds Ash1 mRNA via the adapter protein, She3p.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5273-8.

112.   Shepard, K.A., et al., *Widespread cytoplasmic mRNA transport in yeast: identification of 22 bud-localized transcripts using DNA microarray analysis.* Proc Natl Acad Sci U S A, 2003. **100**(20): p. 11429-34.

113. Oeffinger, M., et al., *Comprehensive analysis of diverse ribonucleoprotein complexes.* Nat Methods, 2007. **4**(11): p. 951-6.

114. Hogan, D.J., et al., *Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system.* PLoS Biol, 2008. **6**(10): p. e255.

115. Manford, A.G., et al., *ER-to-plasma membrane tethering proteins regulate cell signaling and ER morphology.* Dev Cell, 2012. **23**(6): p. 1129-40.

116. Verna, J., et al., *A family of genes required for maintenance of cell wall integrity and for the stress response in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 1997. **94**(25): p. 13804-9.

117. Imazu, H. and H. Sakurai, *Saccharomyces cerevisiae heat shock transcription factor regulates cell wall remodeling in response to heat shock.* Eukaryot Cell, 2005. **4**(6): p. 1050-6.

118. Leon, S., Z. Erpapazoglou, and R. Haguenauer-Tsapis, *Ear1p and Ssh4p are new adaptors of the ubiquitin ligase Rsp5p for cargo ubiquitylation and sorting at multivesicular bodies.* Mol Biol Cell, 2008. **19**(6): p. 2379-88.

119. Herskowitz, I., *Cellular differentiation, cell lineages, and transposable genetic cassettes in yeast.* Curr Top Dev Biol, 1983. **18**: p. 1-14.

120. Nasmyth, K., *Molecular analysis of a cell lineage.* Nature, 1983. **302**(5910): p. 670-6.

121. Strathern, J.N., et al., *Homothallic switching of yeast mating type cassettes is initiated by a double-stranded cut in the MAT locus.* Cell, 1982. **31**(1): p. 183-92.

122. Nasmyth, K.A., *Molecular genetics of yeast mating type.* Annu Rev Genet, 1982. **16**: p. 439-500.

123. Herskowitz, I., *Life cycle of the budding yeast Saccharomyces cerevisiae.* Microbiol Rev, 1988. **52**(4): p. 536-53.

124. Sil, A. and I. Herskowitz, *Identification of asymmetrically localized determinant, Ash1p, required for lineage-specific transcription of the yeast HO gene.* Cell, 1996. **84**(5): p. 711-22.

125. Bobola, N., et al., *Asymmetric accumulation of Ash1p in postanaphase nuclei depends on a myosin and restricts yeast mating-type switching to mother cells.* Cell, 1996. **84**(5): p. 699-709.

126. Jansen, R.P., et al., *Mother cell-specific HO expression in budding yeast depends on the unconventional myosin myo4p and other cytoplasmic proteins.* Cell, 1996. **84**(5): p. 687-97.

127. Haarer, B.K., et al., *Identification of MYO4, a second class V myosin gene in yeast.* J Cell Sci, 1994. **107 (Pt 4)**: p. 1055-64.

128. Zahner, J.E., H.A. Harkins, and J.R. Pringle, *Genetic analysis of the bipolar pattern of bud site selection in the yeast Saccharomyces cerevisiae.* Mol Cell Biol, 1996. **16**(4): p. 1857-70.

129. Pruyne, D., et al., *Stable and dynamic axes of polarity use distinct formin isoforms in budding yeast.* Mol Biol Cell, 2004. **15**(11): p. 4971-89.

130. Toi, H., et al., *She4p/Dim1p interacts with the motor domain of unconventional myosins in the budding yeast, Saccharomyces cerevisiae.* Mol Biol Cell, 2003. **14**(6): p. 2237-49.

131. Wesche, S., M. Arnold, and R.P. Jansen, *The UCS domain protein She4p binds to myosin motor domains and is essential for class I and class V myosin function.* Curr Biol, 2003. **13**(9): p. 715-24.

132. Bohl, F., et al., *She2p, a novel RNA-binding protein tethers ASH1 mRNA to the Myo4p myosin motor via She3p.* EMBO J, 2000. **19**(20): p. 5514-24.

133. Long, R.M., et al., *She2p is a novel RNA-binding protein that recruits the Myo4p-She3p complex to ASH1 mRNA.* EMBO J, 2000. **19**(23): p. 6592-601.

134. Bertrand, E., et al., *Localization of ASH1 mRNA particles in living yeast.* Mol Cell, 1998. **2**(4): p. 437-45.

135. Hasegawa, Y., K. Irie, and A.P. Gerber, *Distinct roles for Khd1p in the localization and expression of bud-localized mRNAs in yeast.* RNA, 2008. **14**(11): p. 2333-47.

136. Irie, K., et al., *The Khd1 protein, which has three KH RNA-binding motifs, is required for proper localization of ASH1 mRNA in yeast.* EMBO J, 2002. **21**(5): p. 1158-67.

137. Gu, W., et al., *A new yeast PUF family protein, Puf6p, represses ASH1 mRNA translation and is required for its localization.* Genes Dev, 2004. **18**(12): p. 1452-65.

138. Deng, Y., R.H. Singer, and W. Gu, *Translation of ASH1 mRNA is repressed by Puf6p-Fun12p/eIF5B interaction and released by CK2 phosphorylation.* Genes Dev, 2008. **22**(8): p. 1037-50.

139. Muller, M., et al., *A cytoplasmic complex mediates specific mRNA recognition and localization in yeast.* PLoS Biol, 2011. **9**(4): p. e1000611.

140. Chartrand, P., et al., *Asymmetric sorting of ash1p in yeast results from inhibition of translation by localization elements in the mRNA.* Mol Cell, 2002. **10**(6): p. 1319-30.

141. Niedner, A., F.T. Edelmann, and D. Niessing, *Of social molecules: The interactive assembly of ASH1 mRNA-transport complexes in yeast.* RNA Biol, 2014. **11**(8): p. 998-1009.

142. Olivier, C., et al., *Identification of a conserved RNA motif essential for She2p recognition and mRNA localization to the yeast bud.* Mol Cell Biol, 2005. **25**(11): p. 4752-66.

143. Jambhekar, A., et al., *Unbiased selection of localization elements reveals cis-acting determinants of mRNA bud localization in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 2005. **102**(50): p. 18005-10.

144. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction.* Nucleic Acids Res, 2003. **31**(13): p. 3406-15.

145. Heym, R.G., et al., *In vitro reconstitution of an mRNA-transport complex reveals mechanisms of assembly and motor activation.* J Cell Biol, 2013. **203**(6): p. 971-84.

146. Shen, Z., A. St-Denis, and P. Chartrand, *Cotranscriptional recruitment of She2p by RNA pol II elongation factor Spt4-Spt5/DSIF promotes mRNA localization to the yeast bud.* Genes Dev, 2010. **24**(17): p. 1914-26.

147. Kruse, C., et al., *Ribonucleoprotein-dependent localization of the yeast class V myosin Myo4p.* J Cell Biol, 2002. **159**(6): p. 971-82.

148. Du, T.G., et al., *Nuclear transit of the RNA-binding protein She2 is required for translational control of localized ASH1 mRNA.* EMBO Rep, 2008. **9**(8): p. 781-7.

149. Shen, Z., et al., *Nuclear shuttling of She2p couples ASH1 mRNA localization to its translational repression by recruiting Loc1p and Puf6p.* Mol Biol Cell, 2009. **20**(8): p. 2265-75.

150. Niessing, D., et al., *She2p is a novel RNA binding protein with a basic helical hairpin motif.* Cell, 2004. **119**(4): p. 491-502.

151. Gonsalvez, G.B., et al., *RNA-protein interactions promote asymmetric sorting of the ASH1 mRNA ribonucleoprotein complex.* RNA, 2003. **9**(11): p. 1383-99.

152. Muller, M., et al., *Formation of She2p tetramers is required for mRNA binding, mRNP assembly, and localization.* RNA, 2009. **15**(11): p. 2002-12.

153. Singh, N., G. Blobel, and H. Shi, *Hooking She3p onto She2p for myosin-mediated cytoplasmic mRNA transport.* Proc Natl Acad Sci U S A, 2015. **112**(1): p. 142-7.

154. Estrada, P., et al., *Myo4p and She3p are required for cortical ER inheritance in Saccharomyces cerevisiae.* J Cell Biol, 2003. **163**(6): p. 1255-66.

155. Heuck, A., et al., *Monomeric myosin V uses two binding regions for the assembly of stable translocation complexes.* Proc Natl Acad Sci U S A, 2007. **104**(50): p. 19778-83.

156. Hodges, A.R., E.B. Krementsova, and K.M. Trybus, *She3p binds to the rod of yeast myosin V and prevents it from dimerizing, forming a single-headed motor complex.* J Biol Chem, 2008. **283**(11): p. 6906-14.

157.   Shi, H., et al., *Structure of a myosin*adaptor complex and pairing by cargo.* Proc Natl Acad Sci U S A, 2014. **111**(12): p. E1082-90.

158.   Landers, S.M., et al., *She3p possesses a novel activity required for ASH1 mRNA localization in Saccharomyces cerevisiae.* Eukaryot Cell, 2009. **8**(7): p. 1072-83.

159.   Sladewski, T.E., et al., *Single-molecule reconstitution of mRNA transport by a class V myosin.* Nat Struct Mol Biol, 2013. **20**(8): p. 952-7.

160.   Dunn, B.D., et al., *Myo4p is a monomeric myosin with motility uniquely adapted to transport mRNA.* J Cell Biol, 2007. **178**(7): p. 1193-206.

161.   Reck-Peterson, S.L., et al., *Class V myosins.* Biochim Biophys Acta, 2000. **1496**(1): p. 36-51.

162.   Trybus, K.M., *Myosin V from head to tail.* Cell Mol Life Sci, 2008. **65**(9): p. 1378-89.

163.   Heuck, A., et al., *The structure of the Myo4p globular tail and its function in ASH1 mRNA localization.* J Cell Biol, 2010. **189**(3): p. 497-510.

164.   Krementsova, E.B., et al., *Two single-headed myosin V motors bound to a tetrameric adapter protein form a processive complex.* J Cell Biol, 2011. **195**(4): p. 631-41.

165.   Sladewski, T.E. and K.M. Trybus, *A single molecule approach to mRNA transport by a class V myosin.* RNA Biol, 2014. **11**(8): p. 986-91.

166.   Long, R.M., et al., *An exclusively nuclear RNA-binding protein affects asymmetric localization of ASH1 mRNA and Ash1p in yeast.* J Cell Biol, 2001. **153**(2): p. 307-18.

167.   Huh, W.K., et al., *Global analysis of protein localization in budding yeast.* Nature, 2003. **425**(6959): p. 686-91.

168.   Niedner, A., et al., *Role of Loc1p in assembly and reorganization of nuclear ASH1 messenger ribonucleoprotein particles in yeast.* Proc Natl Acad Sci U S A, 2013. **110**(52): p. E5049-58.

169.   Paquin, N., et al., *Local activation of yeast ASH1 mRNA translation through phosphorylation of Khd1p by the casein kinase Yck1p.* Mol Cell, 2007. **26**(6): p. 795-809.

170.   Zhao, Y., D.A. Chapman, and I.M. Jones, *Improving baculovirus recombination.* Nucleic Acids Res, 2003. **31**(2): p. E6-6.

171.   Ponchon, L., et al., *A generic protocol for the expression and purification of recombinant RNA in Escherichia coli using a tRNA scaffold.* Nat Protoc, 2009. **4**(6): p. 947-59.

172.   Berger, I., D.J. Fitzgerald, and T.J. Richmond, *Baculovirus expression system for heterologous multiprotein complexes.* Nat Biotechnol, 2004. **22**(12): p. 1583-7.

173.   Sikorski, R.S. and P. Hieter, *A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae.* Genetics, 1989. **122**(1): p. 19-27.

174.   Gietz, R.D. and A. Sugino, *New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites.* Gene, 1988. **74**(2): p. 527-34.

175.   Lescoute, A. and E. Westhof, *Topology of three-way junctions in folded RNAs.* RNA, 2006. **12**(1): p. 83-93.

176.   Inoue, H., H. Nojima, and H. Okayama, *High efficiency transformation of Escherichia coli with plasmids.* Gene, 1990. **96**(1): p. 23-8.

177.   Sambrook, J.a.R., D. W., *Molecular cloning: a laboratory manual*. 2001: Cold Spring Harbor Laboratory Press.

178.   Ho, S.N., et al., *Site-directed mutagenesis by overlap extension using the polymerase chain reaction.* Gene, 1989. **77**(1): p. 51-9.

179.   Studier, F.W., *Protein production by auto-induction in high density shaking cultures.* Protein Expr Purif, 2005. **41**(1): p. 207-34.

180.   Gasteiger, E., et al., *ExPASy: The proteomics server for in-depth protein knowledge and analysis.* Nucleic Acids Res, 2003. **31**(13): p. 3784-8.

181.   Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.

182.   Rasband, W.S. *ImageJ*. 1997-2016; Available from: http://imagej.nih.gov/ij/.

183.   Edman, P., *A method for the determination of amino acid sequence in peptides.* Arch Biochem, 1949. **22**(3): p. 475.

184.   Cherry, J.M., et al., *Saccharomyces Genome Database: the genomics resource of budding yeast.* Nucleic Acids Res, 2012. **40**(Database issue): p. D700-5.

185.   Cole, C., J.D. Barber, and G.J. Barton, *The Jpred 3 secondary structure prediction server.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W197-201.

186.   McGuffin, L.J., K. Bryson, and D.T. Jones, *The PSIPRED protein structure prediction server.* Bioinformatics, 2000. **16**(4): p. 404-5.

187.   Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.

188.   Petoukhov, M.V., et al., *New developments in the program package for small-angle scattering data analysis.* J Appl Crystallogr, 2012. **45**(Pt 2): p. 342-350.

189.   Janowski, R., et al., *Roquin recognizes a non-canonical hexaloop structure in the 3'-UTR of Ox40.* Nat Commun, 2016. **7**: p. 11032.

190.   Goddard, T.D.a.K., D.G., *SPARKY 3*.

191.   Bergfors, T.M., *Protein crystallization: Techniques, Strategies, and Tips: A Laboratory Manual.* 1999: La Jolla, Calif: International University Line.

192.   Dong, A., et al., *In situ proteolysis for protein crystallization and structure determination.* Nat Methods, 2007. **4**(12): p. 1019-21.

193.   Heras, B. and J.L. Martin, *Post-crystallization treatments for improving diffraction quality of protein crystals.* Acta Crystallogr D Biol Crystallogr, 2005. **61**(Pt 9): p. 1173-80.

194.   Lusty, C.J., *A gentle vapor-diffusion technique for cross-linking of protein crystals for cryocrystallography.* J Appl Crystallography, 1999. **32**: p. 106-112.

195.   Winn, M.D., et al., *Overview of the CCP4 suite and current developments.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.

196.   Luscombe, N.M., R.A. Laskowski, and J.M. Thornton, *NUCPLOT: a program to generate schematic diagrams of protein-nucleic acid interactions.* Nucleic Acids Res, 1997. **25**(24): p. 4940-5.

197.   Krementsov, D.N., E.B. Krementsova, and K.M. Trybus, *Myosin V: regulation by calcium, calmodulin, and the tail domain.* J Cell Biol, 2004. **164**(6): p. 877-86.

198.   Muller, M., A. Heuck, and D. Niessing, *Directional mRNA transport in eukaryotes: lessons from yeast.* Cell Mol Life Sci, 2007. **64**(2): p. 171-80.

199.   Lange, S., et al., *Simultaneous transport of different localized mRNA species revealed by live-cell imaging.* Traffic, 2008. **9**(8): p. 1256-67.

200.   Ferre-D'Amare, A.R., K. Zhou, and J.A. Doudna, *A general module for RNA crystallization.* J Mol Biol, 1998. **279**(3): p. 621-31.

201.   Wu, L., et al., *Structural variation and uniformity among tetraloop-receptor interactions and other loop-helix interactions in RNA crystal structures.* PLoS One, 2012. **7**(11): p. e49225.

202.   Coonrod, L.A., J.R. Lohman, and J.A. Berglund, *Utilizing the GAAA tetraloop/receptor to facilitate crystal packing and determination of the structure of a CUG RNA helix.* Biochemistry, 2012. **51**(42): p. 8330-7.

203.   Müller, M., *Characterization of She2p-dependent mRNP assembly in Saccharomyces cerevisiae.*, in *Faculty of Chemistry and Pharmacy*. 2009, LMU Munich.

204.   Kabsch, W., *Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants.* J. Appl. Crystallogr., 1993. **26**: p. 795-800.

205.   McCoy, A.J., et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.

206. Emsley, P., et al., *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.

207. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

208. Morin, A., et al., *Cutting edge: Collaboration gets the most out of software.* eLIFE, 2013. **2**: p. e01456.

209. Vagin, A. and A. Teplyakov, *Molecular replacement with MOLREP.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 22-5.

210. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method.* Acta Crystallogr D Biol Crystallogr, 1997. **53**(Pt 3): p. 240-55.

211. Terwilliger, T.C., *Automated structure solution, density modification and model building.* Acta Cryst. D, 2002. **58**(Pt 11): p. 1937-1940.

212. Heras, B., et al., *Dehydration converts DsbG crystal diffraction from low to high resolution.* Structure, 2003. **11**(2): p. 139-45.

213. Winn, M., M. Isupov, and G.N. Murshudov, *Use of TLS parameters to model anisotropic displacements in macromolecular refinement.* Acta Crystallogr D Biol Crystallogr, 2000. **57**: p. 122-133.

214. Joosten, R.P., et al., *The PDB_REDO server for macromolecular structure model optimization.* IUCrJ, 2014. **1**(Pt 4): p. 213-20.

215. Salemme, F.R., *A free interface diffusion technique for the crystallization of proteins for X-ray crystallography.* Arch Biochem Biophys, 1972. **151**(2): p. 533-9.

216. Beach, D.L., E.D. Salmon, and K. Bloom, *Localization and anchoring of mRNA in budding yeast.* Curr Biol, 1999. **9**(11): p. 569-78.

217. Besse, F., et al., *Drosophila PTB promotes formation of high-order RNP particles and represses oskar translation.* Genes Dev, 2009. **23**(2): p. 195-207.

218. Bullock, S.L., et al., *Guidance of bidirectional motor complexes by mRNA cargoes through control of dynein number and activity.* Curr Biol, 2006. **16**(14): p. 1447-52.

219. Amrute-Nayak, M. and S.L. Bullock, *Single-molecule assays reveal that RNA localization signals regulate dynein-dynactin copy number on individual transcript cargoes.* Nat Cell Biol, 2012. **14**(4): p. 416-23.

220. Lunde, B.M., C. Moore, and G. Varani, *RNA-binding proteins: modular design for efficient function.* Nat Rev Mol Cell Biol, 2007. **8**(6): p. 479-90.

221. Berry, K.E., et al., *Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning.* Structure, 2011. **19**(10): p. 1456-66.

222. Edelmann, F.T., et al., *Molecular architecture and dynamics of ASH1 mRNA recognition by its mRNA-transport complex.* Nat Struct Mol Biol, 2017.

223. Müller, M., et al., *A cytoplasmic complex mediates specific mRNA recognition and localization in yeast.* PLoS Biol, 2011. **9**(4): p. e1000611.

224. Jambhekar, A., et al., *Unbiased selection of localization elements reveals cis-acting determinants of mRNA bud localization in Saccharomyces cerevisiae.* Proc Natl Acad Sci USA, 2005. **102**: p. 18005-18010.

225. Olivier, C., et al., *Identification of a conserved RNA motif essential for She2p recognition and mRNA localization to the yeast bud.* Mol Cell Biol, 2005. **25**(11): p. 4752-4766.

226. McGinnis, J.L., et al., *In-cell SHAPE reveals that free 30S ribosome subunits are in the inactive state.* Proc Natl Acad Sci U S A, 2015. **112**(8): p. 2425-30.

227. Furtig, B., et al., *Multiple conformational states of riboswitches fine-tune gene regulation.* Curr Opin Struct Biol, 2015. **30**: p. 112-24.

228.   Chen, W. and M.J. Moore, *The spliceosome: disorder and dynamics defined.* Curr Opin Struct Biol, 2014. **24**: p. 141-9.

229.   Calabretta, S. and S. Richard, *Emerging Roles of Disordered Sequences in RNA-Binding Proteins.* Trends Biochem Sci, 2015. **40**(11): p. 662-72.

7 Publications

Parts of this work have been published in the following articles:

Edelman, F.T., Schlundt A., Heym R.G., Jenner A., Niedner-Boblenz A., Syed M.I., Paillart J-C., Stehle R., Janowski R., Sattler M., Jansen R-P. & Niessing D. (2017) Molecular architecture and dynamics of *ASH1* mRNA recognition by its mRNA-transport complex. *Nat Struct Mol Biol*, doi: 10.1038/nsmb.3351.

Edelmann, F.T.*, A. Niedner*, and D. Niessing (2015) *ASH1* mRNP-core factors form stable complexes in absence of cargo RNA at physiological conditions. *RNA Biol*, 12(3): p. 233-7.

Edelmann, F.T.*, A. Niedner*, and D. Niessing (2014) Production of pure and functional RNA for *in vitro* reconstitution experiments. *Methods*, 65(3): p. 333-41.

Niedner, A.*, Edelmann F.T.*, and D. Niessing (2014) Of social molecules: The interactive assembly of *ASH1* mRNA-transport complexes in yeast. *RNA Biol*, 11(8): p. 998-1009.

Heym, R.G., Zimmermann, D., Edelmann, F.T., Israel, L., Okten, Z., Kovar, D.R., Niessing, D. (2013) *In vitro* reconstitution of an mRNA-transport complex reveals mechanisms of assembly and motor activation. *J Cell Biol*, 203(6): p. 971-84.

Other publications:

Sun, L.*, Edelmann F.T.*, Kaiser C.J.O.*, Papsdorf K., Gaiser A.M., Richter K. (2012) The lid domain of *Caenorhabditis elegans* Hsc70 influences ATP turnover, cofactor binding and protein folding activity. *PLoS One*, 7(3):e33980.

*Authors contributed equally to this work

Acknowledgements

Hereby I want to thank: